

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

On: Thu Jul 30 09:30:16 1998; MasPar time 4.32 Seconds
278.747 Million cell updates/sec
Bular output not generated.

Title: >US-08-938-548A-11
Description: (1-33) from US08938548A.pap
Perfect Score: 256
Sequence: 1 QPLPCCRQKTCSCRLYELHAGNHAAGILTL 33

Scoring table: PAM 150
Gap 11
Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir56
1:pir1 2:pir2 3:pir3 4:pir4 5:nrl3d

Statistics: Mean 30.310; Variance 51.743; scale 0.586

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	72	28.1	147	2	S24303	hypothetical protein
2	70	27.3	870	2	A41130	dystrophin homolog -
3	69	27.0	93	2	S60079	oct2 protein isoform
4	69	27.0	143	2	A37199	sperm mitochondrial c
5	68	26.6	641	2	JN0800	tetracycline-minocycl
6	67	26.2	299	1	AHRB	ig alpha chain C regi
7	67	26.2	338	2	S09276	ig alpha chain C regi
8	66	25.8	347	2	S11223	UDPglucose 4-epimeras
9	66	25.8	425	2	I56329	gene D3 protein - mou
10	66	25.8	479	2	B69764	transcriptional regul
11	65	25.4	155	2	C64314	hypothetical protein
12	65	25.4	197	2	I46413	keratin KAP5.5 - shee
13	65	25.4	589	2	S33920	beta-fructofuranosida
14	64	25.0	35	5	IOMB	Omega-aga-ivb (nmr, m
15	64	25.0	48	5	IOMA	Omega-aga-ivb (nmr, 2
16	64	25.0	48	5	IAGG	Omega-aga-ivb (nmr, 2
17	64	25.0	48	1	A44664	Omega-aga-ivb (nmr, 2
18	64	25.0	76	2	B54252	Omega-aga-ivb (nmr, 2
19	64	25.0	76	2	A42335	Omega-aga-ivb (nmr, 2
20	64	25.0	76	2	A54252	Omega-aga-ivb (nmr, 2
21	64	25.0	103	2	B64052	cell division protein
22	64	25.0	131	2	A46298	pigment deposition co
23	64	25.0	132	2	I37143	agouti protein precur

24	64	25.0	369	2	B64835	hypothetical protein
25	64	25.0	493	2	S49175	cysteine proteinase (
26	63	24.6	236	2	H69619	hypothetical protein
27	63	24.6	360	2	S09271	ig alpha chain C regi
28	63	24.6	396	2	B57479	amino acid transport
29	63	24.6	477	2	A29978	angiotensin precursor
30	63	24.6	639	2	S13142	tetracycline resistan
31	63	24.6	639	2	A60633	tetracycline resistan
32	63	24.6	639	2	A56779	tetracycline resistan
33	62	24.2	157	2	JC5551	hypothetical 17.9K pr
34	62	24.2	612	2	JH0799	laminin-related prote
35	62	24.2	722	2	B61231	myosin heavy chain, n
36	62	24.2	1464	2	S47555	N-methyl-D-aspartate
37	62	24.2	1464	2	S29159	glutamate receptor, N
38	62	24.2	1464	2	A34274	N-methyl D-aspartate
39	62	24.2	1466	2	A36426	SPA2 protein - yeast
40	61	23.8	76	2	D54252	omega-agatoxin III, 8
41	61	23.8	160	1	WBEC7	gene 34 protein - equ
42	61	23.8	270	2	A42887	urease-associated pro
43	61	23.8	270	2	S32937	ured protein - Klebsi
44	61	23.8	1169	2	S64859	DNA repair protein RA
45	61	23.8	2007	1	B43402	myosin heavy chain-B,

ALIGNMENTS

RESULT 1

ENTRY S24303 #type complete
TITLE hypothetical protein H19-3', - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 10-Sep-1997
ACCESSIONS S24303
REFERENCE S24302
#authors Poirier, F.; Chan, C.T.J.; Timmons, P.M.; Robertson, E.J.; Evans, M.J.; Rigby, P.W.J.
#journal Development (1991) 113:1105-1114
#title The murine H19 gene is activated during embryonic stem cell differentiation in vitro and at the time of implantation in the developing embryo.

#accession S24303
#status preliminary
#molecule_type mRNA
#residues 1-147 #label POI
#cross-references EMBL:X58196; NID:g51131; PID:g51133
SUMMARY #length 147 #molecular-weight 16309 #checksum 6958

Query Match 28.1%; Score 72; DB 2; Length 147;
Best Local Similarity 32.1%; Pred. No. 9.01e-01;
Matches 9; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

Db 66 CCNQHYYTCLRNLIQSGAERTDGVNII 93
||| : : : ||| : : : : :
QY 6 CCRQKTCSCRLYELHAGNHAAGILTL 33

RESULT 2

ENTRY A41130 #type fragment
TITLE dystrophin homolog - Pacific electric ray (fragment)
ORGANISM 300k subsynaptic protein
ALTERNATE_NAMES #formal_name torpado californica #common_name Pacific electric ray
DATE 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 23-Jan-1998
ACCESSIONS A41130
REFERENCE A41130
#authors Readon, J.E.; Lin, H.; Dyer, S.M.; Burden, S.J.
#journal J. Cell Biol. (1991) 115:1069-1076
#title Dystrophin is a component of the subsynaptic membrane.
#cross-references MUID:92064638
#accession A41130
#molecule_type mRNA
#residues 1-870 #label YEA

stabilized by interchain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger oligomers.
#superfamily immunoglobulin C region; immunoglobulin homology duplication; glycoprotein; heterotetramer; immunoglobulin; plasma

CLASSIFICATION

KEYWORDS

FEATURE

86-152 #domain immunoglobulin homology #label IGG1\

189-261 #domain immunoglobulin homology #label IGG2\

38,286 #binding-site carbohydrate (Asn) (covalent) #status predicted

SUMMARY

#length 299 #checksum 2361

Query Match 26.28; Score 67; DB 1; Length 299;

Best Local Similarity 50.08; Pred. No. 5.28e+00;

Matches 7; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Db 44 PFPDCCPANSCTC 57

1:|||||::|:|

2 PLPDCCCRQKTC-SC 14

RESULT

ENTRY 7

TITLE

ORGANISM

DATE

ACCESSIONS

REFERENCE

#authors

#journal

#title

#cross-references MUID:90076124

#accession

#status

#molecule_type DNA

#residues 1-338 #label BUR

CLASSIFICATION

KEYWORDS

FEATURE

SUMMARY

228-300

#domain immunoglobulin homology #label IGG2

#length 338 #checksum 2169

Query Match 26.28; Score 67; DB 2; Length 338;

Best Local Similarity 50.08; Pred. No. 5.28e+00;

Matches 7; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Db 83 PFPDCCPANSCTC 96

1:|||||::|:|

QY 2 PLPDCCCRQKTC-SC 14

RESULT

ENTRY 8

TITLE

ALTERNATE_NAMES

ORGANISM

DATE

ACCESSIONS

REFERENCE

#authors

#journal

#title

#cross-references MUID:90384840

#accession

#status

#molecule_type mRNA

preliminary

#molecule_type mRNA

#residues 1-347 #label ZES

#cross-references EMBL:X53949; NID:g57791; PID:g57792

GENETICS

#gene

CLASSIFICATION

KEYWORDS

FEATURE

SUMMARY

5-343

#domain UDPglucose 4-epimerase homology #label UDP

#length 347 #molecular-weight 38225 #checksum 6947

Query Match 25.88; Score 66; DB 2; Length 347;

Best Local Similarity 53.88; Pred. No. 7.44e+00;

Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 261 CGCRIYNLGTGTG 273

1:|||||::|:|

QY 12 CSCRLYELLHGAG 24

RESULT

ENTRY 9

TITLE

ORGANISM

DATE

ACCESSIONS

REFERENCE

#authors

#journal

#title

#cross-references MUID:93274206

#accession

#status

#molecule_type mRNA

#residues 1-425 #label RES

#cross-references GB:S62227; NID:g385702; PID:g385703

GENETICS

#gene

SUMMARY

D3

#length 425 #molecular-weight 47046 #checksum 8487

Query Match 25.88; Score 66; DB 2; Length 425;

Best Local Similarity 31.08; Pred. No. 7.44e+00;

Matches 9; Conservative 10; Mismatches 9; Indels 1; Gaps 1;

Db 178 PCCCEPTAMCO-SPILHSSSSASSNLS 205

1:|||||::|:|

QY 4 PDCRQKTCRLYELLHGAGNHAAGILT 32

RESULT

ENTRY 10

TITLE

ORGANISM

DATE

ACCESSIONS

REFERENCE

#authors

B69764

#type complete

transcriptional regulator (GntR family) / homolog ycnf -

Bacillus subtilis

#formal_name Bacillus subtilis

05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change

B69764

A69580

Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;

Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.;

Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans,

A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;

Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;

Choi, S.K.; Codani, J.J.; Connerton, I.F.; Cummings, N.J.;

Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoeft, A.;

Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.;

Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita,

M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim,

S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.;

Guilleppi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,

1-23 #domain signal sequence #status experimental #label sig\
 24-589 #product beta-fructofuranosidase #status experimental
 #label MAT\
 36,42,170,188,211,
 254,259,318,322,
 388,463,518,527
 #binding_site carbohydrate (Asn) (covalent) #status
 predicted\
 70-92 #binding_site phosphate (Thr) (covalent) #status
 predicted\
 458,475,490 #binding_site phosphate (Ser) (covalent) #status
 predicted
 SUMMARY #length 589 #molecular-weight 63650 #checksum 1032
 Query Match 25.4%; Score 65; DB 2; Length 589;
 Best Local Similarity 42.1%; Pred. No. 1.05e+01;
 Matches 8; Conservative 6; Mismatches 4; Indels 1; Gaps 1;
 Db 548 RLFDVLNG-GEQAIETL 565
 ||::||:| |::| | |
 15 RLVELLHGAGNHAGI 33

RESULT 14
 ENTRY 1OMB #type complete
 TITLE Omega-aga-ivb (nmr, minimized average structure) -
 funnel-weaving spider (Agelenopsis aperta)
 ORGANISM #formal_name Agelenopsis aperta
 REFERENCE A51323
 #authors Yu, H.; Rosen, M.K.; Schreiber, S.L.
 #submission submitted to the Brookhaven Protein Data Bank, September 1993
 #cross-references PDB:1OMB
 COMMENT Resolution: not applicable
 COMMENT Determination: NMR
 FEATURE
 7-9,31-33,23-24 #region beta sheet\
 10-13 #region turn (type II)\
 1-17 #disulfide_bonds\
 9-22 #disulfide_bonds\
 16-33 #disulfide_bonds\
 24-31 #disulfide_bonds\
 SUMMARY #length 35 #molecular-weight 3804 #checksum 7419
 Query Match 25.0%; Score 64; DB 5; Length 35;
 Best Local Similarity 45.5%; Pred. No. 1.46e+01;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Db 16 CCRGRPCRCGM 26
 ||::||:| |::| | |
 6 CCRQKTCSCRL 16

RESULT 15
 ENTRY 1OMA #type complete
 TITLE Omega-aga-ivb (nmr, 21 structures) - funnel-weaving spider
 (Agelenopsis aperta)
 ORGANISM #formal_name Agelenopsis aperta
 REFERENCE A51322
 #authors Yu, H.; Rosen, M.K.; Schreiber, S.L.
 #submission submitted to the Brookhaven Protein Data Bank, September 1993
 #cross-references PDB:1OMA
 COMMENT Resolution: not applicable
 COMMENT Determination: NMR
 FEATURE
 10-12,34-36,26-27 #region beta sheet\
 13-16 #region turn (type II)\
 4-20 #disulfide_bonds\
 12-25 #disulfide_bonds\
 19-36 #disulfide_bonds\
 27-34 #disulfide_bonds\
 SUMMARY #length 48 #molecular-weight 5281 #checksum 7904
 Query Match 25.0%; Score 64; DB 5; Length 48;
 Best Local Similarity 45.5%; Pred. No. 1.46e+01;

Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
 Db 19 CCRGRPCRCGM 29
 ||::||:| |::| | |
 QY 6 CCRQKTCSCRL 16
 Search completed: Thu Jul 30 09:30:35 1998
 Job time : 19 secs.

MPERCH_PP

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPERCH_PP protein - protein database search, using Smith-Waterman algorithm

on: Thu Jul 30 09:32:10 1998; MasPar time 1.21 Seconds
159.253 Million cell updates/sec
Molecular output not generated.

Title: >US-08-938-548A-11
Description: (1-33) from US08938548A.pap
Perfect Score: 256
Sequence: 1 QPLPCCRQKTCRLYELHAGNHAAGIILTL 33

Scoring table: PAM 150
Gap 11

Searched: 63816 seqs, 5850866 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5-COMB 2:PCPT9_COMB 3:backfiles

Statistics: Mean 21.451; Variance 82.349; scale 0.260

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	66	25.8	39	1	US-08-249- Sequence 40, Applicati	2.23e+01
2	66	25.8	39	1	US-08-036- Sequence 40, Applicati	2.23e+01
3	66	25.8	39	2	PCT-US95-0 Sequence 40, Applicati	2.23e+01
4	66	25.8	39	2	PCT-US94-0 Sequence 40, Applicati	2.23e+01
5	66	25.8	39	1	US-08-469- Sequence 40, Applicati	2.23e+01
6	66	25.8	348	2	PCT-US95-0 Sequence 2, Applicatio	2.23e+01
7	64	25.0	1167	1	US-08-485- Sequence 6, Applicatio	3.29e+01
8	64	25.0	1168	1	US-08-620- Sequence 9, Applicatio	3.29e+01
9	62	24.2	1464	1	US-08-026- Sequence 1, Applicatio	4.83e+01
10	61	23.8	42	1	US-08-487- Sequence 23, Applicati	5.85e+01
11	61	23.8	42	1	US-08-480- Sequence 23, Applicati	5.85e+01
12	61	23.8	42	1	US-08-477- Sequence 23, Applicati	5.85e+01
13	61	23.8	42	1	US-08-137- Sequence 23, Applicati	5.85e+01
14	61	23.8	59	1	US-08-233- Sequence 51, Applicati	5.85e+01
15	60	23.4	222	3	5223425-6 Patent No. 5223425	7.08e+01
16	59	23.4	334	1	US-08-279- Sequence 10, Applicati	7.08e+01
17	59	23.0	283	1	US-08-658- Sequence 2, Applicatio	8.55e+01
18	59	23.0	1019	1	US-08-296- Sequence 4, Applicatio	8.55e+01
19	59	23.0	1083	1	US-08-296- Sequence 2, Applicatio	8.55e+01
20	57	22.3	21	1	US-08-019- Sequence 5, Applicatio	1.24e+02
21	57	22.3	48	1	US-08-379- Sequence 1, Applicatio	1.24e+02
22	57	22.3	724	1	US-08-121- Sequence 62, Applicati	1.24e+02
23	57	22.3	724	2	PCT-US94-1 Sequence 62, Applicati	1.24e+02

24	57	22.3	790	1	US-08-363- Sequence 2, Applicatio	1.24e+02
25	57	22.3	1184	1	US-08-446- Sequence 20, Applicati	1.24e+02
26	57	22.3	1184	1	US-08-446- Sequence 20, Applicati	1.24e+02
27	57	22.3	1187	1	US-08-357- Sequence 8, Applicatio	1.24e+02
28	57	22.3	1187	1	US-08-097- Sequence 13, Applicatio	1.24e+02
29	57	22.3	1187	2	PCT-US95-1 Sequence 8, Applicatio	1.24e+02
30	56	21.9	28	1	US-07-977- Sequence 49, Applicati	1.50e+02
31	56	21.9	37	3	5177308-12 Patent No. 5177308	1.50e+02
32	56	21.9	38	3	5177308-8 Patent No. 5177308	1.50e+02
33	56	21.9	583	3	5475086-4 Patent No. 5475086	1.50e+02
34	56	21.9	584	1	US-08-161- Sequence 1, Applicatio	1.50e+02
35	56	21.9	585	3	5475086-7 Patent No. 5475086	1.50e+02
36	56	21.9	585	1	US-08-485- Sequence 12, Applicati	1.50e+02
37	56	21.9	585	1	US-08-485- Sequence 13, Applicati	1.50e+02
38	56	21.9	585	3	5475086-6 Patent No. 5475086	1.50e+02
39	56	21.9	585	1	US-08-117- Sequence 2, Applicatio	1.50e+02
40	56	21.9	585	1	US-08-485- Sequence 11, Applicatio	1.50e+02
41	56	21.9	1239	1	US-08-026- Sequence 3, Applicatio	1.50e+02
42	56	21.9	2842	1	US-08-289- Sequence 7, Applicatio	1.50e+02
43	56	21.9	2842	1	US-08-452- Sequence 7, Applicatio	1.50e+02
44	56	21.9	2842	1	US-07-741- Sequence 7, Applicatio	1.50e+02
45	56	21.9	2843	1	US-08-452- Sequence 2, Applicatio	1.50e+02

ALIGNMENTS

RESULT 1
ID US-08-249-322A-40 STANDARD; PRT; 39 AA.
XX
AC
XX
XX
DT

Sequence 40, Application US/08249322A

Sequence 40, Application US/08249322A

Patent No. 5716930

GENERAL INFORMATION:

APPLICANT: Goodearl, Andrew; Stroobant, Paul;

APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;

APPLICANT: Chen, Maio Su; Hiles, Ian

TITLE OF INVENTION: Glial Mitogenic Factors, Their

TITLE OF INVENTION: Preparation and Use

NUMBER OF SEQUENCES: 184

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage

COMPUTER: IBM

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/249,322A

FILING DATE: 26-MAY-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/036,555

FILING DATE: 24-MAR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/965,173

FILING DATE: 23-OCT-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/940,389

FILING DATE: 03-SEP-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/907,138

FILING DATE: 30-JUN-1992

PRIOR APPLICATION DATA:

CC FILING DATE: 03-SEP-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/907,138
 CC FILING DATE: 30-JUN-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/863,703
 CC FILING DATE: 03-APRIL-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: U.K. 91 07566.3
 CC FILING DATE: 10-APRIL-1991
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Hanson, Norman D.
 CC REGISTRATION NUMBER: 30,946
 CC REFERENCE/DOCKET NUMBER: LUD 5250.5
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (212) 688-9200
 CC TELEFAX: (212) 838-3884
 CC INFORMATION FOR SEQ ID NO: 40:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 39
 CC TYPE: amino acid
 CC STRANDEDNESS:
 CC TOPOLOGY: linear
 CC SEQUENCE 39 AA; 3760 MW; 6521 CN;
 CC
 CC Query Match 25.8%; Score 66; DB 2; Length 39;
 CC Best Local Similarity 70.0%; Pred. No. 2.23e+01;
 CC Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
 CC
 Db 8 CCR-TTCACR 16
 QY 6 CCRQKTCSCR 15
 CC
 CC RESULT 4
 CC ID PCT-US94-05083C-40 STANDARD; PRT; 39 AA.
 CC XX
 CC AC xxxxxx
 CC XX
 CC DT
 CC XX
 CC DE
 CC XX
 CC Sequence 40, Application PC/TUS9405083C
 CC Sequence 40, Application PC/TUS9405083C
 CC GENERAL INFORMATION:
 CC APPLICANT: Robert Sklar, Mark Marchionni,
 CC APPLICANT: David I. Gwynne
 CC TITLE OF INVENTION: METHODS FOR ALTERING
 CC TITLE OF INVENTION: MUSCLE CONDITION
 CC NUMBER OF SEQUENCES: 185
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Fish & Richardson
 CC STREET: 225 Franklin Street
 CC CITY: Boston
 CC STATE: Massachusetts
 CC ZIP: 02110-2804
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Diskette, 5.25 inch, 360
 CC MEDIUM TYPE: kb storage
 CC COMPUTER: IBM
 CC OPERATING SYSTEM: PC-DOS
 CC SOFTWARE: Wordperfect
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: PCT/US94/05083C
 CC FILING DATE: 06-MAY-94
 CC CLASSIFICATION:
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/209,204
 CC FILING DATE: 08-MAR-94
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/059,022
 CC FILING DATE: 06-MAY-93
 CC ATTORNEY/AGENT INFORMATION:

CC NAME: Clark, Paul T.
 CC REGISTRATION NUMBER: 30,162
 CC REFERENCE/DOCKET NUMBER: 04585/028W01
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 542-5070
 CC TELEFAX: (617) 542-8906
 CC TELEX: 200154
 CC INFORMATION FOR SEQ ID NO: 40:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 39
 CC TYPE: amino acid
 CC STRANDEDNESS:
 CC TOPOLOGY: linear
 CC SEQUENCE 39 AA; 3760 MW; 6521 CN;
 CC
 CC Query Match 25.8%; Score 66; DB 2; Length 39;
 CC Best Local Similarity 70.0%; Pred. No. 2.23e+01;
 CC Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
 CC
 Db 8 CCR-TTCACR 16
 QY 6 CCRQKTCSCR 15
 CC
 CC RESULT 5
 CC ID US-08-469-569-40 STANDARD; PRT; 39 AA.
 CC XX
 CC AC xxxxxx
 CC XX
 CC DT
 CC XX
 CC DE
 CC XX
 CC Sequence 40, Application US/08469569
 CC Sequence 40, Application US/08469569
 CC Patent No. 5606032
 CC GENERAL INFORMATION:
 CC APPLICANT: Goodearl, Andrew; Stroobant, Paul;
 CC APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
 CC APPLICANT: Chen, Maio Su; Hiles, Ian
 CC TITLE OF INVENTION: Glial Mitogenic Factors, Their
 CC TITLE OF INVENTION: Preparation and Use
 CC NUMBER OF SEQUENCES: 184
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Felfe & Lynch
 CC STREET: 805 Third Avenue
 CC CITY: New York City
 CC STATE: New York
 CC COUNTRY: USA
 CC ZIP: 10022
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
 CC COMPUTER: IBM
 CC OPERATING SYSTEM: PC-DOS
 CC SOFTWARE: Wordperfect
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/469,569
 CC FILING DATE: 06-JUN-1995
 CC CLASSIFICATION: 530
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/036,555
 CC FILING DATE: 24-MAR-1993
 CC APPLICATION NUMBER: 07/965,173
 CC FILING DATE: 23-OCT-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/940,389
 CC FILING DATE: 03-SEP-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/907,138
 CC FILING DATE: 30-JUN-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/863,703
 CC FILING DATE: 03-APRIL-1992
 CC PRIOR APPLICATION DATA:

CC REFERENCE/DOCKET NUMBER: MA48DD2.C1
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (904) 375-8100
 CC TELEFAX: (904) 372-5800
 CC INFORMATION FOR SEQ ID NO: 6:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 1167 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 1167 AA; 131657 MW; 7095219 CN;
 SQ
 Query Match 25.0%; Score 64; DB 1; Length 1167;
 Best Local Similarity 45.5%; Pred. No. 3.29e+01;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Db 873 PNCCKPAACQC 883
 I::I: :| |
 4 PDCCRQKTCSC 14
 RESULT 8
 ID US-08-620-717A-9 STANDARD; PRT; 1168 AA.
 XX
 AC xxxxxx
 XX
 DT
 DE
 DE Sequence 9, Application US/08620717A
 XX Sequence 9, Application US/08620717A
 CC Patent No. 5670365
 CC GENERAL INFORMATION:
 CC APPLICANT: Feitelson, Jerald S.
 CC TITLE OF INVENTION: Identification of, and Uses For, Nematocidal
 CC TITLE OF INVENTION: Bacillus thuringiensis Genes, Toxins, and Isolates
 CC NUMBER OF SEQUENCES: 9
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Saliwanchik & Saliwanchik
 CC STREET: 2421 N.W. 41st Street, Suite A-1
 CC CITY: Gainesville
 CC STATE: Florida
 CC COUNTRY: USA
 CC ZIP: 32606
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patentin
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/620,717A
 CC FILING DATE:
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/540,104
 CC FILING DATE: 06-OCT-1995
 CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Saliwanchik, David R.
 CC REGISTRATION NUMBER: 31,794
 CC REFERENCE/DOCKET NUMBER: MA94.C1
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (352) 375-8100
 CC TELEFAX: (352) 372-5800
 CC INFORMATION FOR SEQ ID NO: 9:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 1168 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC ORIGINAL SOURCE:

CC INDIVIDUAL ISOLATE: 167P
 SQ SEQUENCE 1168 AA; 131542 MW; 7175894 CN;
 Query Match 25.0%; Score 64; DB 1; Length 1168;
 Best Local Similarity 45.5%; Pred. No. 3.29e+01;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Db 873 PNCCKPAACQC 883
 I::I: :| |
 4 PDCCRQKTCSC 14
 RESULT 9
 ID US-08-026-138E-1 STANDARD; PRT; 1464 AA.
 XX
 AC xxxxxx
 XX
 DT
 DE
 DE Sequence 1, Application US/08026138E
 XX Sequence 1, Application US/08026138E
 CC Patent No. 5502166
 CC GENERAL INFORMATION:
 CC APPLICANT: Masayoshi MISHINA
 CC TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING THE SAME
 CC NUMBER OF SEQUENCES: 19
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Nishiohata Residence 1-107
 CC STREET: 5214, Nishiohata-machi
 CC CITY: Niigata-shi
 CC STATE: Niigata-ken
 CC COUNTRY: JAPAN
 CC ZIP: 951
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
 CC COMPUTER: IBM Compatible
 CC OPERATING SYSTEM: MS-DOS v.5
 CC SOFTWARE: Word Perfect 5.1
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/026,138E
 CC FILING DATE: 26-FEB-1993
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: JP 39563/1992
 CC FILING DATE: 26-FEB-1992
 CC APPLICATION NUMBER: JP 173155/1992
 CC FILING DATE: 30-JUN-1992
 CC APPLICATION NUMBER: JP 215017/1992
 CC FILING DATE: 12-AUG-1992
 CC APPLICATION NUMBER: JP 303878/1992
 CC FILING DATE: 13-NOV-1992
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Hamburg, C. Bruce
 CC REGISTRATION NUMBER: 22,389
 CC REFERENCE/DOCKET NUMBER: F-4551
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (212) 986-2340
 CC TELEFAX: (212) 953-7733
 CC INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 1464 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single strand
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC ORIGINAL SOURCE:
 CC ORGANISM: mouse
 CC TISSUE TYPE: cerebellum
 CC PUBLICATION INFORMATION:
 CC AUTHORS: Masayoshi MISHINA
 CC TITLE: NOVEL PROTEINS AND GENES CODING THE SAME
 CC RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 to 1464
 CC SEQUENCE 1464 AA; 165489 MW; 11224000 CN;
 SQ

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:29:43 1998; MasPar time 7.52 seconds

67.356 Million cell updates/sec

cellular output not generated.

```

Title: >US-08-938-548A-11
Description: (1-33) from US08938548A.pep
Perfect Score: 256
Sequence: 1 OPLPDCCROKTCSCRLYELLHGAGNI

```

Scoring table: PAM 150
Gap 11

Searched: 124785 seqs, 15338987 residues

Post-processing: Minimum Match 08
Listing first 45

Database:

a-geneseq31-2

1:part1	2:part2	3:part3	4:part4	5:part5	6:part6	7:part7
8:part8	9:part9	10:part10	11:part11	12:part12	13:part13	
14:part14	15:part15	16:part16	17:part17	18:part18		
19:part19	20:part20	21:part21	22:part22	23:part23		
24:part24	25:part25	26:part26	27:part27			

Statistics: Mean 22.991; Variance 84.736; scale 0.271

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB	ID	Description	Pred. No.
1	77	30.1	363	23	W12414	Porcine complement in	6.26e+00
2	75	29.3	373	18	R98461	Murine ICE-ced-3 homo	9.46e+00
3	75	23.3	373	13	R66767	Murine interleukin-1	9.46e+00
4	69	27.0	329	26	W29877	Xenopus phosphatidic acid	3.20e+01
5	66	25.8	102	1	P95679	Xenopus Vgl protein f1	5.80e+01
6	66	25.8	348	20	W101619	Human uridine diphosph	5.80e+01
7	65	25.4	78	5	P20020	Sequence of a foot an	7.06e+01
8	64	25.0	45	23	W10106	Human agouti signall1	8.59e+01
9	64	25.0	45	23	W10105	Murine agouti signall	8.59e+01
10	64	25.0	48	9	A45611	AGI toxin.	8.59e+01
11	64	25.0	48	9	R60293	Calcium channel inhib	8.59e+01
12	64	25.0	48	8	R44209	A. aperta venom fract	8.59e+01
13	64	25.0	130	23	W10102	Human agouti signall1	8.59e+01
14	64	25.0	131	23	W10101	Murine agouti signall	8.59e+01
15	64	25.0	1167	27	W31504	Nematode toxin 167P p	8.59e+01
16	64	25.0	1167	20	W10653	Bacillus thuringiensis	8.59e+01
17	64	25.0	1168	23	W16326	Nematocidal toxin 167	8.59e+01
18	62	24.2	34	10	R55088	Tarantula spider veno	1.27e+02
19	62	24.2	1464	10	R55529	Human NMDA R2A rece	1.27e+02

PN	EP - 48455-A.
PD	31-MAR-1982.
PF	17-SEP-1981.
PR	PR 18-SEP-1980.
PR	PR 2-OCT-1980.
PR	PR 27-NOV-1980.
PR	PR 08-APR-1981.
PR	PR 18-AUG-1981.
	GB-030208.
	GB-034130.
	GB-038147.
	GB-011064.
	GB-025150.

RESULT	9
ID	W10105 standard; protein; 45 AA.
AC	W10105;
DT	17-SEP-1997 (first entry)

PS Example; Page 17: 20pp; English.
 CC The sequence is that of a polypeptide present in fraction K of the
 CC venom of *Agelesopsis aperta*, it blocks calcium channels in cells
 CC of both mammals and invertebrates, partic. those affecting neuronal
 CC and muscle cells. It may be used in the treatment of angina,
 CC hypertension, cardiomyopathies, supraventricular arrhythmia,
 CC oesophageal achalasia, premature labour, and Raynaud's disease.
 CC It may also be of use in the study of cell physiology and in the
 CC control of invertebrate pests. It may be produced synthetically.
 SQ Sequence 48 AA;

Query Match 25.0%; Score 64; DB 8; Length 48;
 Best Local Similarity 45.5%; Pred. No. 8.59e+01;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

DB 19 cccrgprccms 29
 QY 6 CCRQKTCSCRL 16
 ||| : ||| :
 ||| : ||| :
 ||| : ||| :

WU 13
 AC W10102 standard; protein; 130 AA.
 DE 17-SEP-1997 (first entry)
 DE Human agouti signalling protein.
 KW Agouti signalling protein; ASP: depigmenting activity; cosmetic;
 KW hyperpigmentary condition; melasma photoageing spots; solar keratosis;
 KW post-inflammatory hyperpigmentation; wound healing; eumelanogenesis;
 KW vitiligo; leucoderma; albinism; hair greying.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT peptide 1..21
 FT region 32
 FT region 86
 FT region /note= "Signal sequence"
 FT region /note= "Start of basic region"
 FT region /note= "Start of Cysteine-rich motif"
 PN W09700892-A2.
 PD 09-JAN-1997.
 PF 21-JUN-1996; U10695.
 PR 23-JUN-1995; US-000436.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Hearing VJ;
 DR WPI: 97-087323/08.
 PT New agouti signal protein peptide(s) and nucleic acids - used for
 PT altering melanin prodn., for treating e.g. melasma photo-ageing
 PT spots, solar keratosis or vitiligo
 PS Claim 5; Page 8-9; 67pp; English.
 CC The sequences given in W10101-29 are biologically active peptides
 CC and fragments of the agouti signalling protein (ASP) which have
 CC depigmenting activity. These peptides are useful for cosmetic purposes
 CC and for clinical application in the prevention or treatment of various
 CC hyperpigmentary conditions and diseases such as melasma photoageing
 CC spots, solar keratosis, and post-inflammatory hyperpigmentation such
 CC as vitiligo, leucoderma, albinism; hair greying.
 CC occurs at sites of wound healing. They can also be used to provide
 CC enhanced eumelanogenesis for treating eg. vitiligo, leucoderma, some
 CC forms of albinism and hair greying.
 SQ Sequence 131 AA;

Query Match 25.0%; Score 64; DB 23; Length 131;
 Best Local Similarity 41.2%; Pred. No. 8.59e+01;
 Matches 7; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

DB 102 papaccdpcscqrcff 118
 QY 2 PLPDCCRQ-KTCSCLY 17
 ||| : ||| :
 ||| : ||| :
 ||| : ||| :

RESULT 15
 ID W31504 standard; Protein; 1167 AA.
 AC W31504;
 DT 07-APR-1998 (first entry)
 DE Nematode toxin 167p protein.
 KW PCR primer; amplify; nematode toxic protein; *Bacillus thuringiensis*;
 KW delta-endotoxin gene; nematode pest control; *Panagrellus redivivus*;
 KW 167P protein.
 OS *Bacillus thuringiensis*.
 PN W09734926-A2.
 PD 25-SEP-1997.
 PF 21-MAR-1997; U04755.
 PR 21-MAR-1996; US-590554.
 PA (MYCO) MYCOGEN CORP.
 PI Fu J, Narva KE, Payne J;
 DR WPI: 97-480163/44.
 DR N-PSDB: T89185.
 PT *Bacillus thuringiensis* toxin gene - useful in recombinant hosts,
 PT particularly plants for the control of nematodes
 PS Claim 4; Page 35-39; 44pp; English.
 CC This sequence represents the protein encoded by a polynucleotide of the
 CC invention. The polynucleotide of the invention is a sequence from a
 CC *Bacillus thuringiensis* (Bt) isolate selected from PS80J1, PS158D5, a
 CC PS167P, PS169E, PS177F1, PS177G, PS204G4 and PS204G6, that encodes a
 CC toxin active against nematodes. This sequence represents the 167P
 CC protein, and is a delta-endotoxin protein. The polynucleotides and toxins
 CC can be used for the control of nematode pests such as *Panagrellus*
 CC redivivus.
 SQ Sequence 1167 AA;

Query Match 25.0%; Score 64; DB 23; Length 130;
 Best Local Similarity 41.2%; Pred. No. 8.59e+01;
 Matches 7; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

DB 101 papaccdpcscqrcff 117
 QY 2 PLPDCCRQ-KTCSCLY 17
 ||| : ||| :
 ||| : ||| :
 ||| : ||| :

RESULT 14
 ID W10101 standard; protein; 131 AA.
 AC W10101;
 DT 17-SEP-1997 (first entry)
 DE Murine agouti signalling protein.

KW Agouti signalling protein; ASP: depigmenting activity; cosmetic;
 KW hyperpigmentary condition; melasma photoageing spots; solar keratosis;
 KW post-inflammatory hyperpigmentation; wound healing; eumelanogenesis;
 KW vitiligo; leucoderma; albinism; hair greying.
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT peptide 1..21
 FT region 32
 FT region /note= "Signal sequence"
 FT region /note= "Start of basic region"
 FT region /note= "Start of Cysteine-rich motif"
 PN W09700892-A2.
 PD 09-JAN-1997.
 PF 21-JUN-1996; U10695.
 PR 23-JUN-1995; US-000436.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PI Hearing VJ;
 DR WPI: 97-087323/08.
 PT New agouti signal protein peptide(s) and nucleic acids - used for
 PT altering melanin prodn., for treating e.g. melasma photo-ageing
 PT spots, solar keratosis or vitiligo
 PS Claim 5; Page 8-9; 67pp; English.
 CC The sequences given in W10101-29 are biologically active peptides
 CC and fragments of the agouti signalling protein (ASP) which have
 CC depigmenting activity. These peptides are useful for cosmetic purposes
 CC and for clinical application in the prevention or treatment of various
 CC hyperpigmentary conditions and diseases such as melasma photoageing
 CC spots, solar keratosis, and post-inflammatory hyperpigmentation such
 CC as vitiligo, leucoderma, albinism; hair greying.
 CC occurs at sites of wound healing. They can also be used to provide
 CC enhanced eumelanogenesis for treating eg. vitiligo, leucoderma, some
 CC forms of albinism and hair greying.
 SQ Sequence 131 AA;

Query Match 25.0%; Score 64; DB 23; Length 131;
 Best Local Similarity 41.2%; Pred. No. 8.59e+01;
 Matches 7; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

DB 102 papaccdpcscqrcff 118
 QY 2 PLPDCCRQ-KTCSCLY 17
 ||| : ||| :
 ||| : ||| :
 ||| : ||| :

RESULT 15
 ID W31504 standard; Protein; 1167 AA.
 AC W31504;
 DT 07-APR-1998 (first entry)
 DE Nematode toxin 167p protein.
 KW PCR primer; amplify; nematode toxic protein; *Bacillus thuringiensis*;
 KW delta-endotoxin gene; nematode pest control; *Panagrellus redivivus*;
 KW 167P protein.
 OS *Bacillus thuringiensis*.
 PN W09734926-A2.
 PD 25-SEP-1997.
 PF 21-MAR-1997; U04755.
 PR 21-MAR-1996; US-590554.
 PA (MYCO) MYCOGEN CORP.
 PI Fu J, Narva KE, Payne J;
 DR WPI: 97-480163/44.
 DR N-PSDB: T89185.
 PT *Bacillus thuringiensis* toxin gene - useful in recombinant hosts,
 PT particularly plants for the control of nematodes
 PS Claim 4; Page 35-39; 44pp; English.
 CC This sequence represents the protein encoded by a polynucleotide of the
 CC invention. The polynucleotide of the invention is a sequence from a
 CC *Bacillus thuringiensis* (Bt) isolate selected from PS80J1, PS158D5, a
 CC PS167P, PS169E, PS177F1, PS177G, PS204G4 and PS204G6, that encodes a
 CC toxin active against nematodes. This sequence represents the 167P
 CC protein, and is a delta-endotoxin protein. The polynucleotides and toxins
 CC can be used for the control of nematode pests such as *Panagrellus*
 CC redivivus.
 SQ Sequence 1167 AA;

TM

Result No.	Query		Score	Match	Length	DB	ID	Description	Pred. No.
1	75	29.3	373	1	ICER_MOUSE	CASPASE-11 PRECURSOR (5.12e-02		
2	73	28.5	260	1	NMA_HUMAN	PUTATIVE TRANSMEMBRANE	1.20e-01		
3	71	27.7	450	1	PPAW_CAEEL	PUTATIVE ACID PHOSPHAT	2.76e-01		
4	70	27.3	214	1	RCSA_TRYBR	60S RIBOSOMAL PROTEIN	4.17e-01		
5	69	27.0	197	1	M1CS_MOUSE	SPERM MITOCHONDRIAL CA	6.26e-01		
6	69	27.0	461	1	YUL2_CAEEL	PUTATIVE FORKHEAD-RELA	6.26e-01		
7	67	26.2	299	1	ALC_RABIT	IG ALPHA CHAIN C REGIO	1.40e+00		
8	67	26.2	923	1	RE11_SCHPO	MEIOTIC RECOMBINATION	1.40e+00		
9	66	25.8	347	1	GALE_RAT	UDP-GLUCOSE 4-EPIMERAS	2.07e+00		
10	66	25.8	348	1	GALE_HUMAN	UDP-GLUCOSE 4-EPIMERAS	2.07e+00		
11	66	25.8	360	1	PURK_PSEAE	PHOSPHORIBOSYLAMINOIMI	2.07e+00		
12	66	25.8	425	1	IF15_MOUSE	INTERFERON-ACTIVATABLE	2.07e+00		
13	65	25.4	155	1	Y115_METJA	HYPOTHETICAL PROTEIN M	3.06e+00		
14	65	25.4	273	1	MD12_SCHPO	MITOCHONDRIAL INHERITA	3.06e+00		
15	64	25.0	76	1	TX03_AGEAP	OMEGA-AGATOXIN I11A	4.51e+00		
16	64	25.0	83	1	TX4B_AGEAP	OMEGA-AGATOXIN IVB PRE	4.51e+00		
17	64	25.0	103	1	BOLA_HAE1IN	BOLA PROTEIN HOMOLOG	4.51e+00		
18	64	25.0	125	1	AGSW_VULVU	AGOUTI SWITCH PROTEIN	4.51e+00		
19	64	25.0	131	1	AGSW_MOUSE	AGOUTI SWITCH PROTEIN	4.51e+00		
20	64	25.0	132	1	AGSW_HUMAN	AGOUTI SWITCH PROTEIN	4.51e+00		
21	64	25.0	493	1	VPE_VICSA	VACUOLAR PROCESSING EN	4.51e+00		
22	63	24.6	236	1	ECSC_BACSU	PROTEIN ECSC.	6.60e+00		
23	63	24.6	477	1	ANGT_MOUSE	ANGIOTENSINOGEN PRECUR	6.60e+00		

CC STRUCTURE OF THE SPERM MITOCHONDRIA.
CC -1- SUBCELLULAR LOCATION: KERATINOUS MITOCHONDRIAL CAPSULE.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- DEVELOPMENTAL STAGE: LATE MEIOTIC AND EARLY HAPLOID CELLS.

DR EMBL; M88463; G459886; -;
DR EMBL; M29603; G567228; -;
DR PIR; A37199; A37199;
DR HSP; P01058; ITAB.
DR MGD; MGI:96945; MCS.
KW MITOCHONDRION; SELENIUM; SPERM; TESTIS; SPERMATOGENESIS.
FT BINDING 7 7 SELENIUM.
FT BINDING 17 17 SELENIUM.
FT BINDING 34 34 SELENIUM.
SQ SEQUENCE 197 AA; 21015 MW; 4E56990C CRC32;

Query Match 27.0%; Score 69; DB 1; Length 197;
Best Local Similarity 71.4%; Pred. No. 6.26e-01;
Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

123 PLKPPCCPQK-CSC 135
||| ||| ||| |||
2 PL-PDCCRQKTCSC 14

RESULT 6
ID YUL2.CAEEL STANDARD; PRT; 461 AA.
AC Q19802;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PUTATIVE FORKHEAD-RELATED TRANSCRIPTION FACTOR F76A1.2.
GN F76A1.2
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; ACCELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA FULTON L.;
RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- SIMILARITY: CONTAINS A FORK-HEAD DOMAIN.
DR EMBL; U27312; G860690; -;
DR WORMPEP; F26A1.2; CE02683.
DR PROSITE; PS00657; FORK HEAD_1; 1.
DR PROSITE; PS00658; FORK HEAD_2; 1.
DR PROSITE; PS50039; FORK HEAD_3; 1.
KW HYPOTHETICAL PROTEIN; DNA-BINDING; NUCLEAR PROTEIN;
KW TRANSCRIPTION REGULATION.
DR DNA_BIND 171 262 FORK-HEAD.
SQ SEQUENCE 461 AA; 54171 MW; FB2C37B7 CRC32;

Query Match 27.0%; Score 69; DB 1; Length 461;
Best Local Similarity 35.3%; Pred. No. 6.26e-01;
Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 221 RHVLCQLFDVQLQVEG 237
|:|:|:|:|:
QY 8 ROKTCSRLYELLHGAG 24

RESULT 7
ID ALC_RABIT STANDARD; PRT; 299 AA.
AC P01879;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE IG ALPHA CHAIN C REGION (FRAGMENT).
OS ORYCTOLAGUS CUNICULUS (RABBIT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; LAGOMORPHA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 84144059.

RA KNIGHT K.L., MARTENS C.L., STOKLOSA C.M., SCHNEIDERMAN R.D.;
RL NUCLEIC ACIDS RES. 12:1657-1670(1984).
CC -1- FUNCTION: IG ALPHA IS THE MAJOR IMMUNOGLOBULIN CLASS IN BODY
CC SECRETIONS. IT MAY SERVE BOTH TO DEFEND AGAINST LOCAL INFECTION
CC AND TO PREVENT ACCESS OF FOREIGN ANTIGENS TO THE GENERAL
CC IMMUNOLOGIC SYSTEM.
CC -1- THIS IMMUNOGLOBULIN BELONGS TO THE IGA-G SUBCLASS. IT WAS ISOLATED
CC FROM A RABBIT HOMOZYGOUS FOR A2, N80, DEL12,15, F71, G75 HEAVY
CC CHAIN HAPLOTYPE.
DR EMBL; X00353; G1576; -;
DR PIR; A02174; AHRB.
DR HSP; P01857; IPFC.
DR PROSITE; PS00290; IG_MHC; 2.
KW IMMUNOGLOBULIN C REGION.
FT NON_TER 1
SQ SEQUENCE 299 AA; 32256 MW; 594CED7C CRC32;
Query Match 26.2%; Score 67; DB 1; Length 299;
Best Local Similarity 50.0%; Pred. No. 1.40e+00;
Matches 7; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Db 44 PPDCCPANSCTC 57
|:|:|:|:|:
QY 2 PLPDCCRQKTC-SC 14

RESULT 8
ID REIL_SCHPO STANDARD; PRT; 923 AA.
AC Q92380;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE MEIOTIC RECOMBINATION PROTEIN RECL1.
GN RECL1.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97231330.
RA LI Y.F., NUMATA M., WAHLS W.P., SMITH G.R.;
RL MOL. MICROBIOL. 23:869-878(1997).
DR EMBL; U70737; G1619901; -;
KW MEIOSIS.
SQ SEQUENCE 923 AA; 107418 MW; B51C7725 CRC32;

Query Match 26.2%; Score 67; DB 1; Length 923;
Best Local Similarity 35.7%; Pred. No. 1.40e+00;
Matches 10; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Db 232 CCDIMRCICLIVNKLSEKSNQTAIIVL 259
||:|:|:|:|:
QY 6 CCRQKTCRLYELLHGAGNHAAGILTL 33

RESULT 9
ID GALE_RAT STANDARD; PRT; 347 AA.
AC P18645;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE UDP-GLUCOSE 4-EPIMERASE (SC 5.1.3.2) (GALACTONALDENASE) (UDP-
DE GALACTOSE 4-EPIMERASE).
GN GALE.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SKELETAL MUSCLE;
RX MEDLINE; 90384840.
RA ZESCHNIGK M., WILCKEN-BERGEMANN B., STARZINSKI-POWITZ A.;
RL NUCLEIC ACIDS RES. 18:5289-5289(1990).
CC -1- FUNCTION: CATALYZES TWO DISTINCT BUT ANALOGOUS REACTIONS: THE


```

RESULT 13
ID Y115 METJA STANDARD; PRT; 155 AA.
AC Q57579;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE HYPOTHETICAL PROTEIN MJ0115.
GN MJ0115.
OS METHANOCOCUS JANNASCHII.
OC ARCHAEABACTERIA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBEK R., KIRNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GÖGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESSE C.R., VENTER J.C.;
RL SCIENCE 273:1058-1073(1996).
DR EMBL; U67469; G1498882; -
DR TIGR; M70115; -
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 155 AA; 17727 MW; 9D9D61E8 CRC32;

Query Match 25.4%; Score 65; DB 1; Length 155;
Best Local Similarity 53.3%; Pred. No. 3.06e+00;
Matches 8; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Db 72 CCKIKPCPYRDYEL 86
QY 6 CCR-ORTCSRLYEL 19
||: ||: |||
||: ||: |||

RESULT 14
ID MD12-SCHPO STANDARD; PRT; 273 AA.
AC Q92377;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE MITOCHONDRIAL INHERITANCE COMPONENT MD12.
GN MD12.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-972;
RA BERGER K.H., SOGO L.F., YAFFE M.P.;
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SIMILARITY: TO YEAST MD12.
DR EMBL; U64674; G1655884; -
KW MITOCHONDRION.
SQ SEQUENCE 273 AA; 30517 MW; 9024B3CC CRC32;

Query Match 25.4%; Score 65; DB 1; Length 273;
Best Local Similarity 53.8%; Pred. No. 3.06e+00;
Matches 7; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Db 213 LHHGTGEHASSVI 225
QY 19 LHHGAGNHAAGIL 31
||||:||||:
||||:||||:

RESULT 15
ID TX03-AGEAP STANDARD; PRT; 76 AA.
AC P33034;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE OMEGA-AGATOXIN IIIA.

```

```

OS AGELENOPSIS APERTA (FUNNEL-WEB SPIDER).
OC EUKARYOTA; METAZOA; ARTHROPODA; CHELICERATA; ARACHNIDA; ARANEAE.
RN [1]
RP SEQUENCE.
RX TISSUE=VENOM;
RL MEDLINE; 92129351.
RA VENEMA V.J., SWIDEREK K.M., LEE T.D., HATHAWAY G.M., ADAMS M.E.;
RL J. BIOL. CHEM. 267:2610-2615(1992).
CC -!- FUNCTION: OMEGA-AGATOXIN ARE ANTAGONIST OF VOLTAGE-SENSITIVE
CC CALCIUM CHANNELS. THEY BLOCK INSECT NEUROMUSCULAR TRANSMISSION
CC PRESYNAPTICALLY. POTENT BLOCKER OF N- AND L-TYPE CALCIUM CHANNELS.
CC -!- PTM: SIX DISULFIDE BONDS ARE PRESENT (PROBABLE).
DR PIR: A42335; A42335.
KW VENOM; NEUROTOXIN; CALCIUM CHANNEL INHIBITOR; PRESYNAPTIC NEUROTOXIN.
SQ SEQUENCE 76 AA; 8518 MW; C7EA0E12 CRC32;

```

```

Query Match 25.0%; Score 64; DB 1; Length 76;
Best Local Similarity 38.1%; Pred. No. 4.51e+00;
Matches 8; Conservative 6; Mismatches 5; Indels 2; Gaps 2;

```

```

Db 17 QCCRRNGYSCYSLFGYKSG 37
QY 5 DCCRQKT-CSC-RLYELLHGA 23
:||||: ||| | :

```

Search completed: Thu Jul 30 09:31:09 1998
Job time : 16 secs.

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	77	30.1	363	4	002839	PORCINE MEMBRANE COPAC
2	74	28.9	750	3	Q93473	F14B4.1 (FRAGMENT).
3	73	28.5	683	3	Q22187	T05A1.3.
4	73	28.5	1876	3	Q24453	PHOSPHONOSITIDE 3-KIN
5	73	28.5	1876	3	Q01938	PHOSPHONOSITIDE 3-KIN
6	73	28.5	1876	3	Q24209	PHOSPHONOSITIDE 3-KIN
7	72	28.1	147	10	Q61039	3' ORF.
8	72	28.1	815	9	Q33367	DNA GYRASE B SUBUNIT.
9	70	27.3	426	3	Q01969	SIMILAR TO GALACTOKINA
10	70	27.3	883	12	Q91493	STROPHIN (FRAGMENT).
11	68	26.6	641	9	Q48791	TET.
12	68	26.6	646	9	Q48712	TETRACYCLINE RESISTANC
13	68	26.6	1382	12	Q90975	TYROSYNE KINASE.
14	66	25.8	479	9	P94426	HOMOLOGUE OF REGULATOR
15	65	25.4	197	4	Q28584	KAP5.5 KERATIN PROTEIN
16	65	25.4	589	1	Q13388	BETA-D-FRUCTOFURANOSID
17	64	25.0	310	3	Q01473	CSMID Q456.
18	64	25.0	369	9	P75863	FROM BASES 996879 TO 1
19	64	25.0	397	9	Q44258	1-CARBOXY-3-CHLORO-3,4
20	64	25.0	398	3	Q18373	SELF PROTEIN.

[illegible][illegible]

RC STRAIN=168;
RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; D50453; G1805458; -;
DR EMBL; Z99106; E1182356; -;
SQ SEQUENCE 479 AA; 55166 MW; 3C2D1F5A CRC32;

Query Match 25.88; Score 66; DB 9; Length 479;
Best Local Similarity 41.28; Pred. No. 3.99e+00;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 207 RMYQLKNGKQVKTIM 223
QY 15 RLYELLHGAGNHAAGIL 31

RESULT 15
ID Q28584 PRELIMINARY; PRT; 197 AA.
AC Q28584;
DE 01-NOV-1996 (TREMELREL. 01, CREATED)
01-NOV-1996 (TREMELREL. 01, LAST SEQUENCE UPDATE)
01-NOV-1996 (TREMELREL. 01, LAST ANNOTATION UPDATE)
DE KAPS.5 KERATIN PROTEIN (FRAGMENT).
GN KRTAP5.5.
OS OVIS ARIES (SHEEP).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-WOOL FOLLICLE;
RX MEDLINE; 94358466.
RA JENKINS B.J., POWELL B.C.;
RL J. INVEST. DERMATOL. 103:310-317(1994).
DR EMBL; X73435; G313722; -;
KW KERATIN.
FT NON_TER
SQ SEQUENCE 197 AA; 17474 MW; 731C19CA CRC32;

Query Match 25.48; Score 65; DB 4; Length 197;
Best Local Similarity 58.38; Pred. No. 5.80e+00;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 87 VPVCCRPACSC 98
QY 3 LPDCCRQKTCSC 14

Search completed: Thu Jul 30 09:31:53 1998
time : 25 secs.


```

Best Local Similarity 31.8%; Pred.No. 7.27e+00;
Matches 7; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

Db 370 LQEQHLSVLDTLSEERAGVVSM 391
||.:|:||||::||::||
QY 6 LQGRQLRLLQANGNHAAGILTM 27

RESULT      8
ENTRY       S41307      #type complete
TITLE       transcription initiation factor sigma - Streptomyces griseus
ORGANISM    #formal_name Streptomyces griseus
DATE        06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
              08-Sep-1997
ACCESSIONS  S41307
REFERENCE    S41306
#authors     Marcos, A.T.; Diez, B.; Gutierrez, S.; Fernandez, F.J.;
              Velasco, J.; Martin, J.F.
#submission submitted to the EMBL Data Library, December 1993
#description Organization and expression of the hrdB-sprC gene cluster of
              streptomyces griseus encoding a sigma factor protein and a
              serine protease. Role on growth and sporulation of
              streptomyces.
#accesion    S41307      preliminary
#status      ##molecule_type DNA
##residues   1-510 ##label MAR
#cross-references EMBL:X75952; NID:g440164; PID:g581664
GENETICS
#start_codon GTG
CLASSIFICATION #superfamily transcription initiation factor sigma katF;
               transcription initiation factor sigma katF homology
KEYWORDS       DNA binding; sigma factor; transcription initiation
FEATURE
279-505        #domain transcription initiation factor sigma katF
               homology #label KTF
SUMMARY        #length 510 #molecular_weight 55795 #checksum 1415
               34.9%; Score 67; DB 2; Length 510;
Query Match    Best Local Similarity 31.8%; Pred.No. 7.27e+00;
Matches 7; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

Db 438 LQEQHLSVLDTLSEERAGVVSM 459
||.:|:||||::||::||
QY 6 LQGRQLRLLQANGNHAAGILTM 27

RESULT      9
ENTRY       JN0443      #type complete
TITLE       transcription initiation factor sigma homolog hrdB -
              Streptomyces aureofaciens
ORGANISM    #formal_name Streptomyces aureofaciens
DATE        30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
              08-Sep-1997
ACCESSIONS  JN0443
REFERENCE    JN0442
#authors     Kormanec, J.; Farkasovsky, M.; Poutuckova, L.
#journal     Gene (Jr) 122:63-70
#title       Four genes in Streptomyces aureofaciens containing a domain
              characteristic of principal sigma factors.
#accesion    JN0443
#molecule_type DNA
##residues   1-525 ##label KOR
#cross-references GB:M90411; NID:g153305; PID:g153306
GENETICS
#start_codon GTG
CLASSIFICATION #superfamily transcription initiation factor sigma katF;
               transcription initiation factor sigma katF homology
KEYWORDS       DNA binding; sigma factor; transcription initiation
FEATURE
294-520        #domain transcription initiation factor sigma katF
               homology #label KTF
SUMMARY        #length 525 #molecular_weight 57204 #checksum 9676

```

```
##cross-references EMBL:X79980; NID:g510451; PID:g510452
SUMMARY #length 253 #molecular-weight 27305 #checksum 1526

Query Match 34.4%; Score 66; DB 2; Length 253;
Best Local Similarity 37.5%; Pred. No. 1.01e+01;
Matches 9; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Db 77 PGAPALADSVKLLAAPTQVQDI 100
   ||:|:| :||| | | |
Qy 1 PGPPGLQRLQLQANGNHAAGI 24

RESULT 14
ENTRY S19248 #type complete
TITLE RNA-directed DNA polymerase (EC 2.7.7.49), msDNA specific -
ALTERNATE_NAMES Escherichia coli retron Ec107
ORGANISM DNA nucleotidyltransferase (RNA-directed); reverse
#formal_name Escherichia coli retron Ec107
#text_change 20-Feb-1995 #sequence_revision 15-Oct-1996 #text_change
05-Sep-1997
ACCESSIONS S19248
REFERENCE S19248
#authors Herzer, P.J.; Inouye, S.; Inouye, M.
#journal Mol. Microbiol. (1992) 6:345-354
#title Retron Ec107 is inserted into the Escherichia coli genome by
replacing a palindromic 34bp intergenic sequence.
#cross-references MUID:92204001
#accession S19248
#status preliminary
#molecule_type DNA
#residues 1-319 #label HER
#cross-references EMBL:X62563; NID:g42774; PID:g42775
#experimental_source E. coli wild strain

GENETICS
#note insertion site is 82 min of E. coli K12 genetic map
CLASSIFICATION #superfamily reverse transcriptase
KEYWORDS nucleotidyltransferase
SUMMARY #length 319 #molecular-weight 36363 #checksum 2019

Query Match 34.4%; Score 66; DB 1; Length 319;
Best Local Similarity 50.0%; Pred. No. 1.01e+01;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 287 ALQGRINWLLQINPEDEA 304
   :|||: ||| | : |
Qy 5 GLQRLQLQLQANGNHA 22

RESULT 15
ENTRY S19248 #type complete
TITLE adenylate cyclase (EC 4.6.1.1) - fission yeast
ALTERNATE_NAMES (Schizosaccharomyces pombe)
ORGANISM #formal_name Schizosaccharomyces pombe
DATE 23-Mar-1990 #sequence_revision 23-Mar-1990 #text_change
09-Sep-1997
ACCESSIONS A33988
REFERENCE A33988
#authors Young, D.; Riggs, M.; Field, J.; Wojtek, A.; Broek, D.;
Wigler, M.
#journal Proc. Natl. Acad. Sci. U.S.A. (1989) 86:7989-7993
#title The adenylate cyclase gene from Schizosaccharomyces pombe.
#cross-references MUID:90046723
#accession A33988
#molecule_type DNA
#residues 1-1692 #label YOU
#cross-references GB:M26699; NID:g173338; PID:g173339
REFERENCE A33539
#authors Yamawaki-Kataoka, Y.; Tamaoki, T.; Choe, H.R.; Tanaka, H.;
Kataoka, T.
#journal Proc. Natl. Acad. Sci. U.S.A. (1989) 86:5693-5697
#title Adenylate cyclases in yeast: a comparison of the genes from
Schizosaccharomyces pombe and Saccharomyces cerevisiae.
```

```
##cross-references MUID:89345533
#accession A33539 #status preliminary
#molecule_type DNA
#residues 1-1692 #label YAM
#cross-references GB:M24942; NID:g173378; PID:g173379
#note the authors translated the codon TGC for residue 636 as
Ser, and GGC for residue 1243 as Gly
CLASSIFICATION #superfamily leucine-rich alpha-2-glycoprotein repeat
homology; yeast adenylate cyclase catalytic domain homology
phosphorus-oxygen lyase
KEYWORDS
FEATURE 1328-1413 #domain yeast adenylate cyclase catalytic domain
homology #label YACC
SUMMARY #length 1692 #molecular-weight 190332 #checksum 2609

Query Match 33.9%; Score 65; DB 2; Length 1692;
Best Local Similarity 57.1%; Pred. No. 1.39e+01;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 1539 PVQLOGRLRLRIKS 1552
   | ||||| :| :
Qy 3 PPGLQRLQLQLQA 16

Search completed: Thu Jul 30 09:23:43 1998
Job time : 13 secs.
```

W P E R E F H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm

on: Thu Jul 30 09:24:01 1998; MasPar time 4.11 Seconds
164.767 Million cell updates/sec

ular output not generated.

Title: >US-08-938-548A-9
Description: (1-27) from US08938548A.pep
Perfect Score: 192
Sequence: 1 PGPGLGRLQLRLLQANGNHAAGILTM 27

Scoring table: PAM 150
Gap 15

Searched: 59111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 31.311; Variance 47.835; scale 0.655

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	68	35.4	736	1	ALD_MOUSE	ADRENOLEUKODYSTROPHY P	1.44e+00	
2	68	35.4	879	1	YDBH_ECOLI	HYPOTHETICAL 96.8 KD P	1.44e+00	
3	67	34.9	187	1	SODC_HAEPA	SUPEROXIDE DISMUTASE P	2.09e+00	
4	67	34.9	187	1	SODC_HAEIN	SUPEROXIDE DISMUTASE L	2.09e+00	
5	67	34.9	442	1	HRDB_STRCO	RNA POLYMERASE PRINCIP	2.09e+00	
6	67	34.9	1157	1	PEAL_PICPA	PEROXISOME BIOSYNTHESI	2.09e+00	
7	65	33.9	307	1	CC36_CAEEL	CUTICLE COLLAGEN 36.	4.39e+00	
8	65	33.9	1692	1	CYAA_SCHPO	ADENYLATE CYCLASE (EC	4.39e+00	
9	65	33.9	1992	1	TR12_HUMAN	THYROID RECEPTOR INTER	4.39e+00	
10	64	33.3	199	1	IL11_MOUSE	INTERLEUKIN-11 PRECURS	6.32e+00	
11	64	33.3	624	1	SIR_SNP7	SULFITE REDUCTASE (FER	6.32e+00	
12	64	33.3	1477	1	YORL_YEAST	OLIGOMYCIN RESISTANCE	6.32e+00	
13	63	32.8	702	1	YCBY_ECOLI	HYPOTHETICAL 78.9 KD P	9.05e+00	
14	63	32.8	878	1	ECR_DROME	ECYDSONE RECEPTOR (ECD	9.05e+00	
15	62	32.3	256	1	YREC_SNP2	HYPOTHETICAL 28.7 KD P	1.29e+01	
16	62	32.3	380	1	FR12_MOUSE	FACTOR VIII INTRON 22	1.29e+01	
17	62	32.3	400	1	ASL1_SIN3	ARGININOSUCCINATE SYN	1.29e+01	
18	62	32.3	652	1	RPSD_CAUCR	RNA POLYMERASE SIGMA F	1.29e+01	
19	62	32.3	1690	1	CA44_HUMAN	COLLAGEN ALPHA 4(IV) C	1.29e+01	
20	62	32.3	1758	1	CA24_CAEEL	COLLAGEN ALPHA 2(IV) C	1.29e+01	
21	62	32.3	1763	1	CA24_ASCSU	PROCOLLAGEN ALPHA 2(IV	1.29e+01	
22	61	31.8	134	1	Y652_METJA	HYPOTHETICAL PROTEIN M	1.83e+01	
23	61	31.8	180	1	NEF_HV2NZ	NEGATIVE FACTOR (F-PHO	1.83e+01	

24	61	31.8	222	1	BASR_SALTY	TRANSCRIPTIONAL REGULA	1.83e+01
25	61	31.8	240	1	YDGB_ECOLI	HYPOTHETICAL OXIDOREDU	1.83e+01
26	61	31.8	248	1	PSPA_HUMAN	PULMONARY SURFACTANT-A	1.83e+01
27	61	31.8	255	1	YPEI_RHURU	HYPOTHETICAL 28 KD PRO	1.83e+01
28	61	31.8	368	1	COAL_POVLY	BETA-ARRESTIN VPI.	1.83e+01
29	61	31.8	409	1	ARR2_HUMAN	BETA-ARRESTIN 2.	1.83e+01
30	61	31.8	420	1	ARR2_BOVIN	BETA-ARRESTIN 2 (ARRES	1.83e+01
31	61	31.8	429	1	UL88_HCMVA	PROTEIN UL88.	1.83e+01
32	61	31.8	516	1	LEG1_GOSHI	LEGUMIN PRECURSOR (BET	1.83e+01
33	61	31.8	518	1	BAR2_SCHCO	PHEROMONE B ALPHA 2 RE	1.83e+01
34	61	31.8	690	1	VTER_EBV	PROBABLE DNA PACKAGING	1.83e+01
35	60	31.3	199	1	IL11_HUMAN	INTERLEUKIN-11 PRECURS	2.59e+01
36	60	31.3	378	1	CARA_PSEAE	CARBAMOYL-PHOSPHATE SY	2.59e+01
37	60	31.3	400	1	YL37_CAEEL	HYPOTHETICAL 45.5 KD P	2.59e+01
38	60	31.3	410	1	ARR2_RAT	BETA-ARRESTIN 2.	2.59e+01
39	60	31.3	743	1	CA18_MOUSE	COLLAGEN ALPHA 1(VIII)	2.59e+01
40	60	31.3	744	1	CA18_RABIT	COLLAGEN ALPHA 1(VIII)	2.59e+01
41	60	31.3	780	1	ACON_BOVIN	ACONITATE HYDRATASE, M	2.59e+01
42	60	31.3	821	1	ALCR_EMENT	REGULATORY PROTEIN ALC	2.59e+01
43	60	31.3	829	1	PQOF_PSEFL	COENZYME PQQ SYNTHESIS	2.59e+01
44	60	31.3	1806	1	CA1B_HUMAN	COLLAGEN ALPHA 1(XI) C	2.59e+01
45	60	31.3	5217	1	HTS1_COCCA	HC-TOXIN SYNTHETASE (E	2.59e+01

ALIGNMENTS

RESULT 1
ID ALD_MOUSE STANDARD; PRT; 736 AA.
AC P48410;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE ADRENOLEUKODYSTROPHY PROTEIN HOMOLOG (ALDP).
GN ALD OR ALDGH.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-FIBROTIC LIVER;
RX MEDLINE; 95201443.
RA SARDE C.O., THOMAS J., SADOULET H., GARNIER J.M., MANDEL J.L.;
RL MAMM. GENOME 5:810-813(1994).
CC -!- FUNCTION: PROBABLE TRANSPORTER. COULD BE INVOLVED IN THE
IMPORT OF VLCFA-COA SYNTHETASE INTO THE PEROXISOMAL MEMBRANE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PEROXISOMAL
(PROBABLE).
CC -!- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
(ABC TRANSPORTERS). BELONGS TO THE MDR SUBFAMILY.
CC EMBL; Z33637; G520955; -.
DR MGD; MGI:99672; ALDGH.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-BINDING; GLYCOPROTEIN; TRANSMEMBRANE; TRANSPORT; PEROXISOME.
FT TRANSMEM 92 112 POTENTIAL.
FT TRANSMEM 131 151 POTENTIAL.
FT TRANSMEM 238 258 POTENTIAL.
FT TRANSMEM 333 353 POTENTIAL.
FT TRANSMEM 473 493 POTENTIAL.
FT NP_BIND 507 514 ATP (BY SIMILARITY).
SQ SEQUENCE 736 AA; 81858 MW; D373B00E CRC32;

Query Match 35.4%; Score 68; DB 1; Length 736;
Best Local Similarity 44.4%; Pred. No. 1.44e+00;
Matches 12; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Db 704 AGIPKMOGLQELRLQILGEAAAPVQPL 730

Qy 1 PGPPLGRLQLRLLQANGNHAAGILTM 27

RESULT 2
ID YDBH_ECOLI STANDARD; PRT; 879 AA.
AC P52645; P77502; P76855;

GN CYR1.

Query Match 33.3%; Score 64; DB 1; Length 199;
Best Local Similarity 53.3%; Pred. No. 6.32e+00;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 124 PELGALQARLERLR 138
| :||:||||:
QY 1 PGPPGLQRLRLQ 15

RESULT 11
ID SIR_SNP7 STANDARD; PRT; 624 AA.
AC P30008;
DT 01-APR-1993 (REL. 25, CREATED)
DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE SULFITE REDUCTASE (FERREDOXIN) (EC 1.8.7.1).
GN SIR.
OS SYNECHOCOCUS SP. (STRAIN PCC 7942) (ANACYSTIS NIDULANS R2).
OC PROKARYOTA; GRACILICUTES; OXYPHOTOBACTERIA;
OC CYANOBACTERIA (BLUE-GREEN ALGAE); CHROCOCCALES.
[1]
SEQUENCE FROM N.A.
RX MEDLINE; 93349955.
RA GISELMANN G., KLAUSMEIER P., SCHWENN J.D.;
RL BIOCHIM. BIOPHYS. ACTA 1144:102-106(1993).
CC -1- CATALYTIC ACTIVITY: H(2)S + 3 OXIDIZED FERREDOXIN + 3 H(2)O -
SULFITE + 3 REDUCED FERREDOXIN.
CC -1- COFACTOR: THIS ENZYME CONTAINS ONE SIROHEME AND ONE 4FE-4S
IRON-SULFUR CENTER AS PROSTHETIC GROUPS.
CC -1- SUBUNIT: MONOMER.
CC -1- SIMILARITY: THE C-TERMINAL DOMAIN IS A 4FE-4S/SIROHEME DOMAIN
FOUND IN NITRITE REDUCTASES (EC 1.6.6.4 AND EC 1.7.7.1) AND
SULFITE REDUCTASES (EC 1.8.1.2 AND EC 1.8.7.1).
CC EMBL; Z11755; G38930; -.
DR PIR; S19860; RDYCS7.
DR PROSITE; PS00365; NIR_SIR; 1.
KW OXIDOREDUCTASE; HEME; IRON-SULFUR; 4FE-4S.
FT METAL 446 446 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 452 452 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 491 491 IRON-SULFUR (4FE-4S) (POTENTIAL).
FT METAL 495 495 IRON-SULFUR (4FE-4S) AND SIROHEME
(BY SIMILARITY).
SQ SEQUENCE 624 AA; 70031 MW; 9C48C0F0 CRC32;

Query Match 33.3%; Score 64; DB 1; Length 624;
Best Local Similarity 53.3%; Pred. No. 6.32e+00;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

464 PGLLVRRRLLEEG 478
||| |::|||:
QY 4 PGLQRLRLQ 18

RESULT 12
ID YORL_YEAST STANDARD; PRT; 1477 AA.
AC P53049;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE OLIGOMYCIN RESISTANCE ATP-DEPENDENT PERMEASE YORL.
GN YORL OR YGR281W.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
[1]
RX SEQUENCE FROM N.A.
RX MEDLINE; 96069397.
RA KATZMANN D.J., HALLSTROM T.C., VOET M., WYSOCK W., GOLIN J.,
RA VOLCKAERT G., MOYE-ROWLEY W.S.;
RL MOL. CELL. BIOL. 15:6875-6883(1995).
[2]
RX SEQUENCE FROM N.A.
RX STRAIN=S288C / FY1679;
RX MEDLINE; 97245295.

RA VOLCKAERT G., VOET M., ROBBEN J.;
RL YEAST 13:251-259(1997).
CC -1- FUNCTION: REQUIRED FOR OLIGOMYCIN RESISTANCE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
(ABC TRANSPORTERS). BELONGS TO THE MDR SUBFAMILY.
CC EMBL; Z73066; E243252; -.
DR SCD; L0003083; YORL.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW ATP-BINDING; TRANSMEMBRANE; GLYCOPROTEIN; TRANSPORT.
FT TRANSMEM 207 227 POTENTIAL.
FT TRANSMEM 250 270 POTENTIAL.
FT TRANSMEM 329 349 POTENTIAL.
FT TRANSMEM 350 370 POTENTIAL.
FT TRANSMEM 434 454 POTENTIAL.
FT TRANSMEM 479 499 POTENTIAL.
FT TRANSMEM 616 636 POTENTIAL.
FT TRANSMEM 893 913 POTENTIAL.
FT TRANSMEM 941 961 POTENTIAL.
FT TRANSMEM 1028 1048 POTENTIAL.
FT TRANSMEM 1118 1138 POTENTIAL.
FT TRANSMEM 1142 1162 POTENTIAL.
FT NP_BIND 621 628 ATP (POTENTIAL).
FT NP_BIND 1247 1254 ATP (POTENTIAL).
FT CARBOHYD 16 16 POTENTIAL.
FT CARBOHYD 218 218 POTENTIAL.
FT CARBOHYD 295 295 POTENTIAL.
FT CARBOHYD 661 661 POTENTIAL.
FT CARBOHYD 759 759 POTENTIAL.
FT CARBOHYD 799 799 POTENTIAL.
FT CARBOHYD 1345 1345 POTENTIAL.
FT CARBOHYD 1366 1366 POTENTIAL.
SQ SEQUENCE 1477 AA; 166727 MW; 79B302B8 CRC32;

Query Match 33.3%; Score 64; DB 1; Length 1477;
Best Local Similarity 40.0%; Pred. No. 6.32e+00;
Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

Db 499 GRLQSLLEAPDDPNQMIEM 518
|||||:|:|:|:
QY 8 GRLQRLQANGNHAAGILTM 27

RESULT 13
ID YCBY_ECOLI STANDARD; PRT; 702 AA.
AC P75864;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 78.9 KD PROTEIN IN PYRD-PQIA INTERGENIC REGION.
GN YCBY.
OS ESCHERICHIA COLI.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
[1]
RX SEQUENCE FROM N.A.
RX STRAIN-K12 / MG1655;
RA BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
RA SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
CC -1- SIMILARITY: STRONG, TO H.INFLUENZAE HI0116/15.
CC -1- SIMILARITY: TO B.SUBTILIS YWBD.
CC -1- SIMILARITY: BELONGS TO THE UPF0020 FAMILY.
DR EMBL; AE000197; G1787181; -.
DR ECOGENE; EG13717; YCBY.
DR PROSITE; PS01261; UPF0020; 1.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN_1.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 702 AA; 78854 MW; 3218A412 CRC32;

Query Match 32.8%; Score 63; DB 1; Length 702;
Best Local Similarity 61.5%; Pred. No. 9.05e+00;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Result No.	Score	Query		Length	DB	ID	Description	Pred. No.
		Match						
1	75	39.1	1277	10	Q35921		PAR INTERACTING PROTEIN	3.48e-01
2	73	38.0	249	10	P97382		K+ CHANNEL BETA4 SUBUNIT	7.16e-01
3	68	35.4	1174	4	Q95168		TIGHT JUNCTION PROTEIN	4.15e+00
4	67	34.9	178	9	Q31955		YONC PROTEIN.	5.84e+00
5	67	34.9	462	9	P95644		RNA POLYMERASE SIGMA F	5.84e+00
6	67	34.9	510	9	Q59913		RNA POLYMERASE SIGMA F	5.84e+00
7	67	34.9	514	9	P77951		RNA POLYMERASE SIGMA F	5.84e+00
8	67	34.9	525	9	Q59813		RNA POLYMERASE SIGMA F	5.84e+00
9	67	34.9	528	9	Q59814		RNA POLYMERASE SIGMA F	5.84e+00
10	67	34.9	568	3	Q27212		ARTICULIN P60.	5.84e+00
11	67	34.9	608	9	Q45998		DNA POLYMERASE III TAU	5.84e+00
12	66	34.4	130	9	Q60245		HRP6.	8.20e+00
13	66	34.4	150	9	Q27642		DEOXYCYTIDINE-TRIPHOSPH	8.20e+00
14	66	34.4	233	9	Q54224		UNKNOWN GENE.	8.20e+00
15	66	34.4	304	3	Q22732		T24D5.1.	8.20e+00
16	66	34.4	319	9	Q05804		RNA-DIRECTED DNA POLYM	8.20e+00
17	66	34.4	377	9	Q05983		CARA.	8.20e+00
18	66	34.4	1465	3	Q17909		H06001.2.	8.20e+00
19	65	33.9	580	10	Q60811		RETINOID X RECEPTOR IN	1.15e+01
20	65	33.9	1344	10	Q35851		P160 MYB-BINDING PROTEI	1.15e+01

Db 390 LOEQLHSVLDLTLSEAGVVM 411
 || : | : : | : : : : | : : : |
 QY 6 LQGRLLQRLQANGNHAAGILTM 27

RESULT 6
 ID Q59913 PRELIMINARY; PRT; 510 AA.
 AC Q59913;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
 DE RNA POLYMERASE SIGMA FACTOR.
 GN HRDB.
 OS STREPTOMYCES GRISEUS.
 OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; STREPTOMYCETACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-IMRU 3570;
 RV MEDLINE; 95189101.
 RL MARCOS A.T., DIEZ B., GUTIERREZ S., FERNANDEZ F.J., OGUIZA J.A.,
 GENE 153:41-48(1995).
 CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
 CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
 CC THEN IS RELEASED.
 DR EMBL; X75952; G581664; -.
 DR PROSITE; PS00715; SIGMA70_1; 1.
 DR PROSITE; PS00716; SIGMA70_2; 1.
 KW TRANSCRIPTION REGULATION; SIGMA FACTOR; DNA-DIRECTED RNA POLYMERASE;
 KW DNA-BINDING.
 SQ SEQUENCE 510 AA; 55795 MW; C3CB64EB CRC32;

Query Match 34.9%; Score 67; DB 9; Length 510;
 Best Local Similarity 31.8%; Pred. No. 5.84e+00;
 Matches 7; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

Db 438 LOEQLHSVLDLTLSEAGVVM 459
 || : | : : | : : : : | : : : |
 QY 6 LQGRLLQRLQANGNHAAGILTM 27

RESULT 7
 ID P77951 PRELIMINARY; PRT; 514 AA.
 AC P77951;
 DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
 DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
 DE RNA POLYMERASE SIGMA FACTOR.
 GN HRDB.
 OS STREPTOMYCES GRISEUS.
 OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; STREPTOMYCETACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA SHINKAWA H., HATADA Y., OKADA M., KINASHI H., NIMI O.;
 RL J. BIOCHEM. 118:494-499(1995).
 CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
 CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
 CC THEN IS RELEASED.
 DR EMBL; L08071; G1617256; -.
 DR PROSITE; PS00715; SIGMA70_1; 1.
 DR PROSITE; PS00716; SIGMA70_2; 1.
 KW TRANSCRIPTION REGULATION; SIGMA FACTOR; DNA-DIRECTED RNA POLYMERASE;
 KW DNA-BINDING.
 SQ SEQUENCE 514 AA; 56132 MW; 792FDDFF CRC32;

Query Match 34.9%; Score 67; DB 9; Length 514;
 Best Local Similarity 31.8%; Pred. No. 5.84e+00;
 Matches 7; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

Db 442 LOEQLHSVLDLTLSEAGVVM 463
 || : | : : | : : : : | : : : |
 QY 6 LQGRLLQRLQANGNHAAGILTM 27

RESULT 8
 ID Q59813 PRELIMINARY; PRT; 525 AA.
 AC Q59813;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
 DE RNA POLYMERASE SIGMA FACTOR.
 GN HRDB.
 OS STREPTOMYCES AUREOFACIENS.
 OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; STREPTOMYCETACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 93083996.
 RA KORMANEC J., FARKASOVSKY M., POTUCKOVA L.;
 GENE 122:63-70(1992).
 CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
 CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
 CC THEN IS RELEASED.
 DR EMBL; M90411; G153306; -.
 DR PROSITE; PS00715; SIGMA70_1; 1.
 DR PROSITE; PS00716; SIGMA70_2; 1.
 KW TRANSCRIPTION REGULATION; SIGMA FACTOR; DNA-DIRECTED RNA POLYMERASE;
 KW DNA-BINDING.
 SQ SEQUENCE 525 AA; 57204 MW; 7B7689F1 CRC32;

Query Match 34.9%; Score 67; DB 9; Length 525;
 Best Local Similarity 31.8%; Pred. No. 5.84e+00;
 Matches 7; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

Db 453 LOEQLHSVLDLTLSEAGVVM 474
 || : | : : | : : : : | : : : |
 QY 6 LQGRLLQRLQANGNHAAGILTM 27

RESULT 9
 ID Q59814 PRELIMINARY; PRT; 528 AA.
 AC Q59814;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
 DE RNA POLYMERASE SIGMA FACTOR.
 GN HRDC.
 OS STREPTOMYCES AUREOFACIENS.
 OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; STREPTOMYCETACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 93083996.
 RA KORMANEC J., FARKASOVSKY M., POTUCKOVA L.;
 GENE 122:63-70(1992).
 CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
 CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
 CC THEN IS RELEASED.
 DR EMBL; M90412; G153309; -.
 DR PROSITE; PS00715; SIGMA70_1; 1.
 DR PROSITE; PS00716; SIGMA70_2; 1.
 KW TRANSCRIPTION REGULATION; SIGMA FACTOR; DNA-DIRECTED RNA POLYMERASE;
 KW DNA-BINDING.
 SQ SEQUENCE 528 AA; 57598 MW; CE3227B6 CRC32;

Query Match 34.9%; Score 67; DB 9; Length 528;
 Best Local Similarity 36.4%; Pred. No. 5.84e+00;
 Matches 8; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Db 456 LOEQLHSVLDLTLSEAGVVM 477
 || : | : : | : : : : | : : : |
 QY 6 LQGRLLQRLQANGNHAAGILTM 27

RESULT 10
 ID Q27212 PRELIMINARY; PRT; 568 AA.
 AC Q27212;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)

```

RESULT 15
ID Q22732 PRELIMINARY; PRT; 304 AA.
AC Q22732;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE T24DS.1.
OS CAENORHABDITIS ELEGANS.
OC EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.
RN [1]
RP SEQUENCE FROM N.A.
RA WILKINSON J.;
RN [2]
RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
  BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J.,
  COULSON A., CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A.,
  FULTON L., GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M.,
  RA JOHNSTON L., JONES M., KERSHAW J., KIRSTEN J., LAISTER N.,
  RA LATREILLE P., LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B.,
  RA O'CALLAGHAN M., PARSONS J., PERCY C., RIFKEN L., ROOPRA A.,
  SAUNDERS D., SHOWNKEEN R., SHALDON N., SMITH A., SONNHAMMER E.,
  RA STADEN R., SULSTON J., THIERRY-MIEG J., THOMAS K., VAUDIN M.,
  RA VAUGHAN K., WATERSTON R., WATSON A., WEINSTOCK L.,
  RA WILKINSON-SPROAT J., WOHLDMAN P.;
  RL NATURE 368:32-38(1994).
DR EMBL; Z68012; G1070097; -
SQ SEQUENCE 304 AA; 34775 MW; C4AALEFB CRC32;

Query Match 34.4%; Score 66; DB 3; Length 304;
Best Local Similarity 40.9%; Pred. No. 8.20e+00;
Matches 9; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

Db 238 LELRGOLLGNDHVAHTLYTL 259
|: ||: ||: ||: ||: |
QY 6 LQGRQLRLQANGNHAAGILTM 27

```

Search completed: Thu Jul 30 09:24:41 1998
Job time : 13 secs.

WQREH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Jul 30 09:10:49 1998; MasPar time 3.13 Seconds
Linear output not generated. 137.296 Million cell updates/sec

Title: >US-08-938-548A-4
Description: (1-28) from US08938548A.pap
Perfect Score: 196
Sequence: 1 RSGPPGLQRLQLLQASGNHAGILTM 28

Scoring table: PAM 150
Gap 15

Searched: 124785 seqs, 15338987 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq31-2
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27

Statistics: Mean 21.977; Variance 87.361; scale 0.252

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	68	34.7	716	R99737	Retinoid X receptor 1	4.32e+01
2	67	34.2	1477	W10424	Saccharomyces cerevisiae	5.22e+01
3	67	34.2	1477	R67691	S. cerevisiae scdur2R	5.22e+01
4	67	34.2	1841	W22605	Tyrosine kinase OR	5.22e+01
5	67	34.2	2192	W21732	LexA/NuMA fusion prot	5.22e+01
6	67	34.2	2272	W21731	GAL4/HA/NuMA fusion p	5.22e+01
7	67	34.2	4630	W19629	Streptomyces venezuelae	5.22e+01
8	66	33.7	10	R20235	p33* N-terminal (2)	6.30e+01
9	63	32.1	32	R00579	New polypeptide based	1.10e+02
10	63	32.1	248	P60442	Plasmid PASpcq-SV(10)	1.10e+02
11	63	32.1	248	P80694	Sequence deduced from	1.10e+02
12	63	32.1	248	P80694	35kd pulmonary surfac	1.10e+02
13	63	32.1	248	P82980	Sequence deduced from	1.10e+02
14	63	32.1	248	P60665	Sequence of human alv	1.10e+02
15	63	32.1	248	P60665	Genomic sequence of h	1.10e+02
16	63	32.1	248	R04216	Human 32K ASP encoded	1.10e+02
17	63	32.1	248	P06662	35kd pulmonary surfac	1.10e+02
18	63	32.1	248	R04215	Human 32K ASP encoded	1.10e+02
19	63	32.1	248	P60441	Plasmid PASpc-SV(10)	1.10e+02

20	63	32.1	248	2	R05091	Vector PSP 35K-1A-10	1.10e+02
21	63	32.1	248	1	R06331	Human alveolar surfac	1.10e+02
22	63	32.1	271	1	R04217	Human 32K ASP encoded	1.10e+02
23	63	32.1	271	3	P60661	Genomic sequence of h	1.10e+02
24	63	32.1	271	1	R04212	Human 32K alveolar su	1.10e+02
25	63	32.1	878	6	R32889	DHR23alpha protein.	1.10e+02
26	63	32.1	878	3	R13793	Ecdysone receptor.	1.10e+02
27	62	31.6	176	23	W13056	HIV-2 provirus-encode	1.33e+02
28	62	31.6	510	9	R49835	Thermus aquaticus hea	1.33e+02
29	62	31.6	2509	24	W32881	Protein (OA-519) cros	1.33e+02
30	60	30.6	169	8	R43261	Human adipogenesis in	1.91e+02
31	60	30.6	177	8	R43262	Human adipogenesis in	1.91e+02
32	60	30.6	178	18	W02202	Human interleukin-11	1.91e+02
33	60	30.6	178	14	R75337	Human interleukin-11	1.91e+02
34	60	30.6	199	4	R24436	Sequence of a cytokin	1.91e+02
35	60	30.6	199	10	R50176	Human interleukin-11	1.91e+02
36	60	30.6	199	8	R43260	Human adipogenesis in	1.91e+02
37	60	30.6	274	13	R68001	Orotidine-5'-monophos	1.91e+02
38	60	30.6	296	13	R76812	Thioredoxin-IL-11 fus	1.91e+02
39	60	30.6	296	9	R45916	E.coli thioredoxin-hu	1.91e+02
40	60	30.6	296	5	R26213	Fusion protein of IL-	1.91e+02
41	60	30.6	486	7	R35231	Rat ALF.	1.91e+02
42	60	30.6	528	19	W05831	M. tuberculosis RNA p	1.91e+02
43	60	30.6	530	14	R76480	Virulence-associated	1.91e+02
44	60	30.6	3724	25	W22608	Platenolide synthase	1.91e+02
45	60	30.6	3724	25	W23718	Platenolide synthase	1.91e+02

ALIGNMENTS

RESULT 1
ID R99737 standard; Protein; 716 AA.
AC R99737;
DT 27-SEP-1996 (first entry)
DE Retinoid X receptor interacting protein R1P110.
KW Retinoid X receptor interacting protein; RXR; RIP; R1P110.
OS Mus sp.
PN W09621677-A1.
PF 18-JUL-1996.
PD 08-DEC-1995; U16311.
PR 13-JAN-1995; US-372652.
PA (GENO) GEN HOSPITAL CORP.
PI Choi H, Moore D, Seol W;
DR WPI: 96-342241/34.
DR N-PSDB; T31030.
PT Retinoid X receptor (RXR) interacting protein (RIP) - useful to
modulate or mediate RXR function, anti-RIP antibodies can be used to
determine RIP subcellular distribution patterns
PS Claim 2: Page 48-50; 90pp; English.
CC Mouse retinoid X receptor (RXR) interacting protein R1P110 (R99737)
is a candidate transcriptional co-activator. It was identified using
an in vivo interaction trap system for the isolation of proteins that
physically interact with RXRs, esp. with the ligand binding domain of
human RXR alpha. Recombinant R1P110 can be obt'd using a cDNA clone
(T31930) obt'd. from a mouse liver library. RIPS (see also R99735-36
and R99738-39) can be used to modulate or mediate RXR function, and
may be used therapeutically or to raise antibodies.
SQ Sequence 716 AA;

Query Match 34.7%; Score 68; DB 18; Length 716;
Best Local Similarity 44.4%; Pred.No. 4.32e+01;
Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
Db 621 qgrllsllqsehrtrtg 638
||||| :|:|:|:
QY 8 QGRLLQLLQASGNHAGI 25

RESULT 2
ID W10424 standard; Protein; 1477 AA.
AC W10424;
DT 18-AUG-1997 (first entry)
DE Saccharomyces cerevisiae aureobasidin resistance protien scdur2.

LD	R02235	standard; Ficolin; 10 AM.
AC	R20235;	
AC	24-APR-1992	(first entry)
DE	"p33" N-terminal	(2).
DE	MAP: lymphocyte; IL-2; CTL;	polymorphism.
KW	Homo sapiens.	
OS	Homo sapiens.	
FS	key	Location/Qualifiers
FT	misc_difference	6
FT	/note=	'polymorphism - see CC"
PN	W09200329-A.	
PD	03-JAN-1992.	
PF	27-JUN-1991;	U04588.
PR	27-JUN-1990;	US-544862.
PA	(BIOG-) BIOGEN INC.	

PA (CALB-) Calif Biotechn Inc.
 PI Schilling JW, White RT, Cordell B, Benson BJ;
 DR WPI: 88-124493/18.
 PT Recombinant alveolar surfactant protein - used for treating
 PT respiratory distress syndrome and related diseases e.g. pneumonia
 PT and bronchitis.
 PS Disclosure: pp: English.
 CC The sequence was deduced from clone pHS10-5, isolated from a
 CC human lung cDNA library. The protein is part of the alveolar
 CC surfactant protein, high mol. wt, hydrophilic 32K gp. The protein
 CC differs at position 50 from a previously published (W08603408)
 CC sequence, and differs at several places from two other sequences
 CC determined by others. It is believed that the 32K human ASP
 CC protein may be encoded by multiple genes. The recombinant protein
 CC can be used for the treatment of respiratory disorders.
 CC See also P82977-80, and P82982.
 SQ Sequence 248 AA;

Query Match 32.1%; Score 63; DB 1; Length 248;
 Best Local Similarity 56.3%; Pred. No. 1.10e+02;

Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 95 gppglpahldeeqat 110
 ||||| :||: |||:
 QY 3 GPPGLQRLRLQAS 18

RESULT 12

ID P70663 standard; Protein; 248 AA.
 AC P70663;
 DT 29-APR-1991 (first entry)
 DE 35kd pulmonary surfactant protein.
 KW Hyaline membrane disease; respiratory distress syndrome; RDS.
 OS Homo sapiens.
 PN W08702037-A.
 PD 09-APR-1987.
 PF 26-SEP-1986; U02034.
 PR 26-SEP-1985; US-781130.
 PR 15-AUG-1986; US-897183.
 PA (GENE-) Genetics Institute Inc.
 PA (BRIG-) Brigham and Women's Hospital.
 PI Tausch HW, Cobs KA, Steinbrink DR, Floros J, Phelps DS;
 DR WPI: 87-108682/15.
 DR N-PSDB; N71010.
 PT Pulmonary surfactant proteins - used for treating Hyaline Membrane
 PT Disease or Respiratory Distress Syndrome.
 PS Claim 1; Page 34A-B; 50pp; English.
 CC Gene product may be used in treatment of Hyaline Membrane Disease
 CC and Respiratory Distress Syndrome (RDS) in both premature infants
 CC and adults eg. cardio-pulmonary operations. The protein products
 CC may also be used to raise diagnostic antibodies.
 SQ Sequence 248 AA;

Query Match 32.1%; Score 63; DB 2; Length 248;
 Best Local Similarity 56.3%; Pred. No. 1.10e+02;

Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 95 gppglpahldeeqat 110
 ||||| :||: |||:
 QY 3 GPPGLQRLRLQAS 18

RESULT 13

ID P82980 standard; protein; 248 AA.
 AC P82980;
 DT 04-DEC-1990 (first entry)
 DE Sequence deduced from pHS10-4, encoding human 32K ASP.
 KW Alveolar surfactant protein; ASP; respiratory distress syndrome;
 KW pneumonia; bronchitis; 32K.
 OS Homo sapiens.
 PN W08805820-A.
 PD 11-AUG-1986.
 PF 15-JAN-1987; U00092.

PR 29-JAN-1987; US-008453.
 PA (CALB-) Calif Biotechn Inc.
 PI Schilling JW, White RT, Cordell B, Benson BJ;
 DR WPI: 88-124493/18.
 PT Recombinant alveolar surfactant protein - used for treating
 PT respiratory distress syndrome and related diseases e.g. pneumonia
 PT and bronchitis.
 PS Disclosure: pp: English.
 CC The sequence was deduced from clone pHS10-4, isolated from a
 CC human lung cDNA library. The protein is part of the alveolar
 CC surfactant protein, high mol. wt, hydrophilic 32K gp. The protein
 CC differs at 7 positions from a previously published (W08603408)
 CC sequence, and also at several places from two other sequences
 CC determined by others. It is believed that the 32K human ASP
 CC protein may be encoded by multiple genes. The recombinant protein
 CC can be used for the treatment of respiratory disorders.
 CC See also P82977-79, P82982 and P80694.
 SQ Sequence 248 AA;

Query Match 32.1%; Score 63; DB 1; Length 248;
 Best Local Similarity 56.3%; Pred. No. 1.10e+02;

Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 95 gppglpahldeeqat 110
 ||||| :||: |||:
 QY 3 GPPGLQRLRLQAS 18

RESULT 14

ID P60665 standard; Protein; 248 AA.
 AC P60665;
 DT 31-JUL-1991 (first entry)
 DE Sequence of human alveolar surfactant protein (hASP)
 DE on pMT(E):HS and pASPC-SV(10)
 KW Regulatable expression system.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..20
 FT protein 21..248
 PN W08604920-A.
 PD 28-AUG-1986.
 PF 11-FEB-1986; U00296.
 PR 13-FEB-1985; US-701296.
 PR 25-NOV-1985; US-801674.
 PA (BIOT-) BIOTECHN RES PARTNE.
 PA (KUSH-) KUSHNER P J.
 PI Kushner PJ, Cofer CL, Friedman J, Talmadge KD;
 DR WPI: 86-238888/36.
 DR N-PSDB; N60571.
 PT Regulatable expression systems - contg. human metallo:thionein-II
 PT promoter
 PS Example; Fig 5; 94pp; English.
 CC A regulatable expression system for a coding sequence is claimed.
 CC The system can process genomic as well as intronless DNA.
 SQ Sequence 248 AA;

Query Match 32.1%; Score 63; DB 3; Length 248;
 Best Local Similarity 56.3%; Pred. No. 1.10e+02;

Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 95 gppglpahldeeqat 110
 ||||| :||: |||:
 QY 3 GPPGLQRLRLQAS 18

RESULT 15

ID P60666 standard; Protein; 248 AA.
 AC P60666;
 DT 31-JUL-1991 (first entry)
 DE Genomic sequence of human alveolar surfactant protein (hASP)
 DE encoded by genomic DNA, used to obtain pASPCq-SV(10)
 KW Regulatable expression system.
 OS Homo sapiens.

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```

MFsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Jul 30 09:13:08 1998; MasPar time 1.15 Seconds
142.938 Million cell updates/sec
cellular output not generated.

```

```

Title:
Description: (1-28) from US08938548A.pep
Perfect Score: 196
Sequence: 1 RSGPGLGRLRLQLQASGNHAAIGILTM 28

```

Scoring table: PAM 150
Gap 15

Searched: 63816 seqs, 5850866 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5 COMB 2:PCT9 COMB 3:backfiles

Statistics: Mean 20.276; Variance 81.503; scale 0.249

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query		Length	DB	ID	Description	Pred. No.
	Score	Match					
1	68	34.7	716	2	PCT-US95-1	Sequence 4, Applicatio	1.33e+01
2	66	33.7	10	2	PCT-US91-0	Sequence 3, Applicatio	1.95e+01
3	62	31.6	176	2	PCT-US96-1	Sequence 6, Applicatio	4.17e+01
4	62	31.6	2509	1	US-08-469-	Sequence 10, Applicati	4.17e+01
5	60	30.6	199	1	US-08-792-	Sequence 8, Applicatio	6.07e+01
6	60	30.6	199	1	US-07-949-	Sequence 4, Applicatio	6.07e+01
7	60	30.6	199	1	US-08-115-	Sequence 2, Applicatio	6.07e+01
8	60	30.6	199	2	PCT-US93-0	Sequence 2, Applicatio	6.07e+01
9	60	30.6	199	1	US-08-017-	Sequence 4, Applicatio	6.07e+01
10	60	30.6	199	1	US-07-941-	Sequence 2, Applicatio	6.07e+01
11	60	30.6	296	1	US-08-115-	Sequence 4, Applicatio	6.07e+01
12	60	30.6	296	1	US-07-745-	Sequence 14, Applicati	6.07e+01
13	60	30.6	296	1	US-07-941-	Sequence 4, Applicatio	6.07e+01
14	60	30.6	296	1	US-08-165-	Sequence 14, Applicati	6.07e+01
15	60	30.6	296	1	US-07-921-	Sequence 14, Applicati	6.07e+01
16	60	30.6	296	2	PCT-US93-0	Sequence 4, Applicatio	6.07e+01
17	60	30.6	296	2	PCT-US94-1	Sequence 14, Applicati	6.07e+01
18	60	30.6	1271	2	PCT-US94-0	Sequence 2, Applicatio	6.07e+01
19	59	30.1	829	1	US-08-446-	Sequence 2, Applicatio	7.31e+01
20	59	30.1	829	1	US-08-220-	Sequence 2, Applicatio	7.31e+01
21	59	30.1	829	1	US-08-445-	Sequence 2, Applicatio	7.31e+01
22	59	30.1	829	1	US-07-670-	Sequence 2, Applicatio	7.31e+01
23	59	30.6	829	1	US-07-776-	Sequence 25, Applicati	8.79e+01

CC	CITY: Spring House	
CC	STATE: Pennsylvania	
CC	COUNTRY: USA	
CC	ZIP: 19477	
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE: Floppy disk	
CC	COMPUTER: IBM PC compatible	
CC	OPERATING SYSTEM: PC-DOS/MS-DOS	
CC	SOFTWARE: Patent in Release #1.0, Version #1.25	
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER: US/07/941,372	
CC	FILING DATE: 19920902	
CC	CLASSIFICATION: 424	
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME: Bak, Mary E.	
CC	REGISTRATION NUMBER: 31,215	
CC	REFERENCE/DOCKET NUMBER: INDUS1	
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE: (215) 540-9206	
CC	TELEFAX: (215) 540-5818	
CC	INFORMATION FOR SEQ ID NO: 2:	
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH: 199 amino acids	
CC	TYPE: AMINO ACID	
CC	TOPOLOGY: linear	
CC	MOLECULE TYPE: protein	
CC	SEQUENCE 199 AA: 21429 MW; 188641 CN;	
SQ	Query Match 30.6%; Score 60; DB 1; Length 199;	
	Best Local Similarity 70.0%; Pred. No. 6.07e+01;	
	Matches 7; Conservative 3; Mismatches 0; Indels	
Db	129 LQARLDRLRL 138	
Qy	11:11:111:	
	7 LQGRLQRLLQ 16	
RESULT	11	
ID	US-08-115-680-4	STANDARD; PRT; 296 AA.
XX	XXXXXX	
AC		
XX		
DT		
XX		
DE	Sequence 4, Application US/08115680	
XX		
CC	Sequence 4, Application US/08115680	
CC	Patent No. 5437863	
CC	GENERAL INFORMATION:	
CC	APPLICANT: Williams, David A.	
CC	APPLICANT: Clark, Steven C.	
CC	TITLE OF INVENTION: Method of Treating Cell Damage or	
CC	TITLE OF INVENTION: Depletion	
CC	NUMBER OF SEQUENCES: 4	
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE: Howson and Howson	
CC	STREET: Spring House Corporate Cntr, P.O. Box 457	
CC	CITY: Spring House	
CC	STATE: Pennsylvania	
CC	COUNTRY: USA	
CC	ZIP: 19477	
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE: Floppy disk	
CC	COMPUTER: IBM PC compatible	
CC	OPERATING SYSTEM: PC-DOS/MS-DOS	
CC	SOFTWARE: Patent in Release #1.0, Version #1.25	
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER: US/08/115,680	
CC	FILING DATE:	
CC	CLASSIFICATION: 514	
CC	PRIOR APPLICATION DATA:	
CC	APPLICATION NUMBER: US 07/941,372	
CC	FILING DATE: 02-SEP-1992	

DT XX DE XX XX
Sequence 14, Application US/08165301A
Sequence 14, Application US/08165301A
Patent No. 5646016
GENERAL INFORMATION:
CC APPLICANT: McCoy, John
CC APPLICANT: DiBlasio-Smith, Elizabeth
CC APPLICANT: Grant, Kathleen
CC APPLICANT: Lavallie, Edward R.
CC TITLE OF INVENTION: PEPTIDE AND PROTEIN FUSIONS TO
CC TITLE OF INVENTION: THIOREDIXIN, THIOREDIXIN-LIKE MOLECULES, AND MODIFIED
CC TITLE OF INVENTION: THIOREDIXIN-LIKE MOLECULES
CC NUMBER OF SEQUENCES: 29
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Genetics Institute, Inc.
CC STREET: 87 Cambridgepark Drive
CC CITY: Cambridge
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02140
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/165,301A
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Melinert, M. C.
CC REGISTRATION NUMBER: 33,544
CC REFERENCE/DOCKET NUMBER: GI 5188D
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 876-1170
CC TELEFAX: (617) 876-5851
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 296 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 296 AA; 31769 MW; 414378 CN;
Query Match 30.6%; Score 60; DB 1; Length 296;
Best Local Similarity 70.0%; Pred. No. 6.07e+01;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
226 LQARLDRLLR 235
||:|:|:|:
QY 7 LQGRLQRLIQ 16
RESULT 15
ID US-07-921-848-14 STANDARD; PRT; 296 AA.
XX AC xxxxxx
XX DT
XX DE
Sequence 14, Application US/07921848
Sequence 14, Application US/07921848
Patent No. 5292646
GENERAL INFORMATION:
CC APPLICANT: McCoy, John
CC APPLICANT: Lavallie, Edward
CC TITLE OF INVENTION: Peptide and Protein Fusions To
CC TITLE OF INVENTION: Thioredoxin and Thioredoxin-Like Molecules
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Genetics Institute, Inc.
CC STREET: 87 Cambridgepark Drive
CC CITY: Cambridge
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02140
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/921,848
CC FILING DATE: 19920728
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/652,531
CC FILING DATE: 06-FEB-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/745,382
CC FILING DATE: 14-AUG-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Cseri, Luann
CC REGISTRATION NUMBER: 31,822
CC REFERENCE/DOCKET NUMBER: GI5188A
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 876-1170
CC TELEFAX: (617) 876-5851
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 296 amino acids
CC TYPE: AMINO ACID
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 296 AA; 31769 MW; 414378 CN;
Query Match 30.6%; Score 60; DB 1; Length 296;
Best Local Similarity 70.0%; Pred. No. 6.07e+01;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
226 LQARLDRLLR 235
||:|:|:|:
QY 7 LQGRLQRLIQ 16
Search completed: Thu Jul 30 09:13:14 1998
Job time : 6 secs.

WIREH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:08:48 1998; MasPar time 5.39 Seconds
257.883 Million cell updates/sec

Similar output not generated.

Title: >US-08-938-548A-3
Description: (1-33) from US08938548A.pep
Perfect Score: 256
Sequence: 1 QPLPDCRCQKTCSCRYELHAGNHAAGILTL 33

Scoring table: PAM 150
Gap 11

Searched: 140542 seqs, 42109429 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrembl5
1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phase 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

Statistics: Mean 30.295; Variance 46.908; scale 0.646

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	ID	Description	Pred. No.
1	77	30.1	002839	PORCINE MEMBRANE COFAC	5.22e-02
2	74	28.9	093473	FL1B4.1 (FRAGMENT).	1.77e-01
3	73	28.5	022187	TO5A1.3.	2.65e-01
4	73	28.5	1876	PHOSPHONOSITIDE 3-KIN	2.65e-01
5	73	28.5	001938	PHOSPHONOSITIDE 3-KIN	2.65e-01
6	73	28.5	1876	PHOSPHONOSITIDE 3-KIN	2.65e-01
7	72	28.1	061639	3' ORF.	3.94e-01
8	72	28.1	033367	DNA GYRASE B SUBUNIT.	3.94e-01
9	70	27.3	001969	SIMILAR TO GALACTOKINA	8.65e-01
10	70	27.3	091493	DYSTROPHIN (FRAGMENT).	8.65e-01
11	68	26.6	048791	TET.	1.87e+00
12	68	26.6	048712	TETRACYCLINE RESISTANC	1.87e+00
13	68	26.6	1382	TYROSINE KINASE.	1.87e+00
14	66	25.8	090975	HOMOLOGUE OF REGULATOR	3.99e+00
15	65	25.4	028584	KAP5.5 KERATIN PROTEIN	5.80e+00
16	65	25.4	013388	BETA-D-FRUCTOFURANOSID	8.38e+00
17	64	25.0	001473	COSMID C04E6.	8.38e+00
18	64	25.0	369	FROM BASES 996879 TO 1	8.38e+00
19	64	25.0	397	1-CARBOXY-3-CHLORO-3,4	8.38e+00
20	64	25.0	018373	SELD PROTEIN.	8.38e+00

21	64	25.0	398	3	018597	SELENOPHOSPHATE SYNTHET	8.38e+00
22	63	24.6	146	9	005606	REPRESSOR/INDUCER PROT	1.21e+01
23	63	24.6	188	3	018238	COSMID C27A2.	1.21e+01
24	63	24.6	227	8	004393	RIBONUCLEASE.	1.21e+01
25	63	24.6	366	3	022627	F21B10.6.	1.21e+01
26	63	24.6	405	3	019671	F21C3.1.	1.21e+01
27	63	24.6	466	8	039135	AMINO ACID TRANSPORTER	1.21e+01
28	63	24.6	639	9	053770	TETRACYCLINE RESISTANC	1.21e+01
29	63	24.6	639	11	067709	ORF11.	1.21e+01
30	63	24.6	639	9	057224	TETM GENE.	1.21e+01
31	63	24.6	639	9	047810	LET 858.	1.21e+01
32	63	24.6	897	3	017336	SIMILARITY TO EGF-LIKE	1.21e+01
33	63	24.6	2946	3	018857	NONMUSCLE MYOSIN HEAVY	1.73e+01
34	62	24.2	99	4	027991	NONMUSCLE MYOSIN HEAVY	1.73e+01
35	62	24.2	106	2	012989	NONMUSCLE MYOSIN HEAVY	1.73e+01
36	62	24.2	109	4	027990	NONMUSCLE MYOSIN HEAVY	1.73e+01
37	62	24.2	157	9	033700	DNA FOR DNAB, COMPLETE	1.73e+01
38	62	24.2	272	10	062707	NONMUSCLE MYOSIN HEAVY	1.73e+01
39	62	24.2	282	10	062706	NONMUSCLE MYOSIN HEAVY	1.73e+01
40	62	24.2	304	3	022732	T24D5.1.	1.73e+01
41	62	24.2	1464	2	012879	N-METHYL-D-ASPARTATE R	1.73e+01
42	62	24.2	1464	10	063728	N-METHYL-D-ASPARTATE R	1.73e+01
43	62	24.2	1464	10	008948	N-METHYL-D-ASPARTATE R	1.73e+01
44	61	23.8	282	12	091373	TRKB (FRAGMENT).	2.47e+01
45	61	23.8	2007	12	002015	MYOSIN HEAVY CHAIN, NO	2.47e+01

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	363 AA.
ID	002839			
AC	002839			
DC	01-JUL-1997 (TREMBLREL. 04, CREATED)			
DT	01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)			
DT	01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)			
DE	PORCINE MEMBRANE COFACTOR PROTEIN.			
OS	SUS SCROFA (PIG).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; ARTIODACTYLA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE; 97343414.			
RA	TOYOMURA K., FUJIMURA T., MURAKAMI H., NATSUME T., SHIGEHISA T.,			
RA	INOUE N., TAKEDA J., KINOSHITA T.;			
RL	INT. IMMUNOL. 9:869-876(1997).			
DR	EMBL; D70897; G1018989; -			
KW	MEMBRANE.			
SQ	SEQUENCE 363 AA; 39692 MW; AD14F57A CRC32;			

Query Match 30.1%; Score 77; DB 4; Length 363;
Best Local Similarity 39.3%; Pred. No. 5.22e-02;
Matches 11; Conservative 7; Mismatches 9; Indels 1; Gaps 1;

Db	98	PLQECACRRKACS-NLPDPLNGQVSPNG 124
QY	2	PLPDCRCQKTCSCRYELHAGNHAAG 29

RESULT	2	PRELIMINARY;	PRT;	750 AA.
ID	093473			
AC	093473			

DT	01-FEB-1997 (TREMBLREL. 02, CREATED)			
DT	01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)			
DT	01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)			
DE	FL1B4.1 (FRAGMENT).			
OS	CAENORHABDITIS ELEGANS.			
OC	EUKARYOTA; METAZOA; ACCELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	SIMS M.;			
RL	SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.			
RP	SEQUENCE FROM N.A.			

[illegible]

RC STRAIN-168;
RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;
RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; D50453; G1805458; -
DR EMBL; Z99106; E1182356; -
SQ SEQUENCE 479 AA; 55166 MW; 3C2D1F5A CRC32;

Query Match 25.8%; Score 66; DB 9; Length 479;
Best Local Similarity 41.2%; Pred. No. 3.99e+00;
Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 207 RMYQLLNAGKQVKTIM 223
QY 15 RLYELLHGAGNHAAGIL 31
|:|:| |::: |:
|:|:| |::: |:

RESULT 15
ID Q28584 PRELIMINARY; PRT; 197 AA.
AC Q28584;
DT 01-NOV-1996 (TREMELREL. 01, CREATED)
01-NOV-1996 (TREMELREL. 01, LAST SEQUENCE UPDATE)
01-NOV-1996 (TREMELREL. 01, LAST ANNOTATION UPDATE)
DE KAP5.5 KERATIN PROTEIN (FRAGMENT).
GN KRTAP5.5.
OS OVIS ARIES (SHEEP).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-WOOL FOLLICLE;
RX MEDLINE; 94358466.
RA JENKINS B.J.; POWELL B.C.;
RL J. INVEST. DERMATOL. 103:310-317(1994).
DR EMBL; X73435; G313722; -
KW KERATIN.
FT NON_TER
SQ SEQUENCE 197 AA; 17474 MW; 731C19CA CRC32;

Query Match 25.4%; Score 65; DB 4; Length 197;
Best Local Similarity 58.3%; Pred. No. 5.80e+00;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 87 VPVCCRPACSC 98
QY 3 LPDCCRQKTCSC 14
|:|:| |::: |:
|:|:| |::: |:

Search completed: Thu Jul 30 09:09:09 1998
time : 21 secs.

WQSRFH
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:08:19 1998; MasPar time 3.18 Seconds
Molecular output not generated. 260.375 Million cell updates/sec

Title: >US-08-938-548A-3
Description: (1-33) from US08938548A.pep
Perfect Score: 256
Sequence: 1 QPLPCCRQKTCRLYELHAGNHAAGIULTL 33

Scoring table: PAM 150
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 31.693; Variance 44.992; scale 0.704

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	75	29.3	373	1	ICEB_MOUSE CASPASE-11 PRECURSOR (5.12e-02
2	73	28.5	260	1	NNA_HUMAN PUTATIVE TRANSMEMBRANE	1.20e-01
3	71	27.7	450	1	PPAW_CAEEL PUTATIVE ACID PHOSPHAT	2.76e-01
4	70	27.3	214	1	R10A_TRYBR 60S RIBOSOMAL PROTEIN	4.17e-01
5	69	27.0	197	1	MCS_MOUSE SPERM MITOCHONDRIAL CA	6.26e-01
6	69	27.0	461	1	YUL2_CAEEL PUTATIVE FORKHEAD-RELA	6.26e-01
7	67	26.2	299	1	ALC_RABIT IG ALPHA CHAIN C REGIO	1.40e+00
8	67	26.2	923	1	REIL_SCHPO MEIOTIC RECOMBINATION	1.40e+00
9	66	25.8	347	1	GALE_RAT UDP-GLUCOSE 4-EPIMERAS	2.07e+00
10	66	25.8	348	1	GALE_HUMAN UDP-GLUCOSE 4-EPIMERAS	2.07e+00
11	66	25.8	360	1	PURK_PSEAP PHOSPHORIBOSYLAMINOIM	2.07e+00
12	66	25.8	425	1	IFIS_MOUSE INTERFERON-ACTIVATABLE	2.07e+00
13	65	25.4	155	1	Y115_METJA MITOCHONDRIAL PROTEIN M	3.06e+00
14	65	25.4	273	1	MD12_SCHPO OMEGA-AGATOXIN IIIA.	4.51e+00
15	64	25.0	76	1	TXO3_AGEAP OMEGA-AGATOXIN IVB PRE	4.51e+00
16	64	25.0	83	1	BOLA_HAEIN BOLA PROTEIN HOMOLOG.	4.51e+00
17	64	25.0	103	1	AGSW_VULVU AGOUTI SWITCH PROTEIN	4.51e+00
18	64	25.0	125	1	AGSW_VULVU AGOUTI SWITCH PROTEIN	4.51e+00
19	64	25.0	131	1	AGSW_HUMAN AGOUTI SWITCH PROTEIN	4.51e+00
20	64	25.0	132	1	AGSW_HUMAN AGOUTI SWITCH PROTEIN	4.51e+00
21	64	25.0	493	1	VPE_VICSA VACUOLAR PROCESSING EN	4.51e+00
22	63	24.6	236	1	ECSC_BACSU PROTEIN ECSC.	6.60e+00
23	63	24.6	477	1	ANGT_MOUSE ANGIOTENSINOGEN PRECUR	6.60e+00

24	63	24.6	639	1	TET9_ENTFA	TETRACYCLINE RESISTANC	6.60e+00
25	62	24.2	612	1	UNC6_CAEEL	UNC-6 PROTEIN PRECURSO	9.61e+00
26	62	24.2	1464	1	NME1_RAT	GLUTAMATE (NMDA) RECEP	9.61e+00
27	62	24.2	1464	1	NME1_MOUSE	GLUTAMATE (NMDA) RECEP	9.61e+00
28	62	24.2	1466	1	SPA2_YEAST	SPA2 PROTEIN.	9.61e+00
29	62	24.2	1976	1	MYSO_HUMAN	MYOSIN HEAVY CHAIN, NO	9.61e+00
30	61	23.8	160	1	VG34_HSVB	GENE 34 PROTEIN.	1.39e+01
31	61	23.8	270	1	URED_KLEAE	UREASE OPERON URED PRO	1.39e+01
32	61	23.8	270	1	URED_KLEPN	UREASE OPERON URED PRO	1.39e+01
33	61	23.8	406	1	Y129_HUMAN	HYPOTHETICAL PROTEIN K	1.39e+01
34	61	23.8	442	1	TBB_TRYBR	TUBULIN BETA CHAIN.	1.39e+01
35	61	23.8	483	1	PRPD_SALTY	PRPD PROTEIN.	1.39e+01
36	61	23.8	595	1	BETP_CORGL	GLYCINE BETAINNE TRANSP	1.39e+01
37	61	23.8	1169	1	RAD5_YEAST	DNA REPAIR PROTEIN RAD	1.39e+01
38	60	23.4	334	1	YHO5_YEAST	HYPOTHETICAL 37.9 KD P	2.01e+01
39	60	23.4	371	1	NTF6_TOBAC	MITOGEN-ACTIVATED PROT	2.01e+01
40	60	23.4	471	1	YMT1_CAEEL	HYPOTHETICAL 51.7 KD P	2.01e+01
41	60	23.4	639	1	YAF3_SCHPO	TETRACYCLINE RESISTANC	2.01e+01
42	60	23.4	1039	1	DPOL_ADE07	DNA POLYMERASE (EC 2.7	2.01e+01
43	60	23.4	1122	1	YIC6_YEAST	HYPOTHETICAL 133.0 KD	2.01e+01
44	60	23.4	1150	1	THYG_BOVIN	THYROGLOBULIN PRECURSO	2.01e+01
45	60	23.4	2769	1			

ALIGNMENTS

RESULT	ID	ICEB_MOUSE	STANDARD;	PRT;	373 AA.
AC	P70343;	O08735;			
DT	01-NOV-1997	(REL. 35, CREATED)			
DT	01-NOV-1997	(REL. 35, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997	(REL. 35, LAST ANNOTATION UPDATE)			
DE	CASPASE-11 PRECURSOR (EC 3.4.22.-) (ICH-3 PROTEASE).				
GN	CASP11 OR ICH3 OR CASPL.				
OS	MUS MUSCULUS (MOUSE).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; RODENTIA.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-CS7BL/6 X CBA; TISSUE-THYMUS;				
RX	MEDLINE; 96355393.				
RA	WANG S., MIYURA M., JUNG Y.K., ZHU H., GAGLIARDINI V., SHI L.,				
RA	GREENBERG A.H., YUAN J.,				
RL	J. BIOL. CHEM. 271:20580-20587(1996).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-C3H/AN;				
RX	MEDLINE; 97190206.				
RA	VAN DE CRAEN M., VANDENABEELE P., DECLERCO W., VAN DEN BRANDE I.,				
RA	VAN LOO G., MOLEMAN F., SCHOTTE P., VAN CRIEINGE W., BEYAERT R.,				
RA	FIERS W.;				
RL	FEBS LETT. 403:61-69(1997).				
CC	-1- FUNCTION: INVOLVED IN THE ACTIVATION CASCADE OF CASPASES				
CC	RESPONSIBLE FOR APOPTOSIS EXECUTION (BY SIMILARITY).				
CC	-1- SUBUNIT: HETERODIMER OF TWO SUBUNITS (POTENTIAL).				
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C14; ALSO KNOWN AS THE				
CC	CASPASE FAMILY.				
DR	EMBL; U59463; G1575318; -				
DR	EMBL; Y13089; E315510; -				
DR	MGI; 107700; CASPL.				
DR	PROSITE; PS01121; CASPASE_HIS; 1.				
DR	PROSITE; PS01122; CASPASE_CYS; 1.				
KW	HYDROLASE; THIOL PROTEASE; APOPTOSIS; ZYMOGEN.				
FT	PROPEP 1 79				
FT	CHAIN 80 266				
FT	PROPEP 267 285				
FT	CHAIN 286 373				
FT	ACT_SITE 206 254				
FT	ACT_SITE 254 254				
FT	CONFLICT 152 152				
SQ	SEQUENCE 373 AA; 42756 MW; 70DBA63 CRC32;				
	Query Match	29.3%;	Score 75;	DB 1;	Length 373;

CC STRUCTURE OF THE SPERM MITOCHONDRIA.
 CC -1- SUBCELLULAR LOCATION: KERATINOUS MITOCHONDRIAL CAPSULE.
 CC -1- TISSUE SPECIFICITY: TESTIS.
 CC -1- DEVELOPMENTAL STAGE: LATE MEIOTIC AND EARLY HAPLOID CELLS.

DR EMBL; M88463; G459886; -.
 DR EMBL; M29603; G567228; -.
 DR PIR; A37199; A37199.
 DR HSP; P01058; 1TAB.
 DR MGD; MGI:96945; MCS.
 KW MITOCHONDRION; SELENIUM; SPERM; TESTIS; SPERMATOGENESIS.
 FT BINDING 7 SELENIUM.
 FT BINDING 17 SELENIUM.
 FT BINDING 34 SELENIUM.
 SQ SEQUENCE 197 AA; 21015 MW; 4E56990C CRC32;

Query Match 27.0%; Score 69; DB 1; Length 197;
 Best Local Similarity 71.4%; Pred. No. 6.26e-01;
 Matches 10; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Db 123 PLKPPCCPOK-CSC 135
 2 PL-PDCCRCORTCSC 14

RESULT 6
 ID YUL2 CAEEL STANDARD; PRT; 461 AA.

AC Q19802;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE PUTATIVE FORKHEAD-RELATED TRANSCRIPTION FACTOR F26A1.2.
 GN F26A1.2
 OS CAENORHABDITIS ELEGANS.
 CC EUKARYOTA; METAZOA; ACCELLOMATES; NEMATODA; SECERNENTEIA; RHABDITIDA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA FULTON L.;
 RL SUBMITTED (JUN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
 CC -1- SIMILARITY: CONTAINS A FORK-HEAD DOMAIN.
 DR EMBL; U27312; G860690; -.
 DR WORMPEP; F26A1.2; CE02683.
 DR PROSITE; PS00657; FORK_HEAD_1; 1.
 DR PROSITE; PS00658; FORK_HEAD_2; 1.
 DR PROSITE; PS50039; FORK_HEAD_3; 1.
 KW HYPOTHETICAL PROTEIN; DNA-BINDING; NUCLEAR PROTEIN;
 KW TRANSCRIPTION REGULATION.
 DNA_BIND 171 262 FORK-HEAD.
 SEQUENCE 461 AA; 54171 MW; FB2C37B7 CRC32;

Query Match 27.0%; Score 69; DB 1; Length 461;
 Best Local Similarity 35.3%; Pred. No. 6.26e-01;
 Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 221 RHVLCQQLFDVLOVEG 237
 8 RQKTCRLYELLHGAG 24

RESULT 7
 ID ALC_RABIT STANDARD; PRT; 299 AA.

AC P01879;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
 DE IG ALPHA CHAIN C REGION (FRAGMENT).
 OS ORYCTOLAGUS CUNICULUS (RABBIT).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; LAGOMORPHA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 84144059.

RA KNIGHT K.L., MARTENS C.L., STOKLOSA C.M., SCHNEIDERMAN R.D.;
 RL NUCLEIC ACIDS RES. 12:1657-1670(1984).
 CC -1- FUNCTION: IG ALPHA IS THE MAJOR IMMUNOGLOBULIN CLASS IN BODY
 CC SECRETIONS. IT MAY SERVE BOTH TO DEFEND AGAINST LOCAL INFECTION
 CC AND TO PREVENT ACCESS OF FOREIGN ANTIGENS TO THE GENERAL
 CC IMMUNOLOGIC SYSTEM.
 CC -1- THIS IMMUNOGLOBULIN BELONGS TO THE IGA-G SUBCLASS. IT WAS ISOLATED
 CC FROM A RABBIT HOMOZYGOUS FOR A2, N80, DE12,15, F71, G75 HEAVY
 CC CHAIN HAPLOTYPE.
 DR EMBL; X00353; G1576; -.
 DR PIR; A02174; AHRB.
 DR HSP; P01857; 1PFC.
 DR PROSITE; PS00290; IG_MHC; 2.
 KW IMMUNOGLOBULIN C REGION.
 FT NON_TER 1
 SQ SEQUENCE 299 AA; 32256 MW; 594CED7C CRC32;

Query Match 26.2%; Score 67; DB 1; Length 299;
 Best Local Similarity 50.0%; Pred. No. 1.40e+00;
 Matches 7; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Db 44 PFPDCCPANSCTC 57
 2 PLPDCCRCQKTC-SC 14

RESULT 8
 ID REIL_SCHPO STANDARD; PRT; 923 AA.

AC Q92380;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE MEIOTIC RECOMBINATION PROTEIN RECL1.
 GN RECL1.
 OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
 CC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 97231330.
 RA LI Y.F., NUMATA M., WAHLS W.P., SMITH G.R.;
 RL MOL. MICROBIOL. 23:869-878(1997).
 DR EMBL; U70737; G1619901; -.
 KW MEIOSIS.
 SQ SEQUENCE 923 AA; 107418 MW; B51C7725 CRC32;

Query Match 26.2%; Score 67; DB 1; Length 923;
 Best Local Similarity 35.7%; Pred. No. 1.40e+00;
 Matches 10; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

Db 232 CCDIMRCLCLIVNKLSEKSNQTAELVL 259
 6 CCRQKTCRLYELLHGAGNHAAGILT 33

RESULT 9
 ID GALE_RAT STANDARD; PRT; 347 AA.

AC P18645;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE UDP-GLUCOSE 4-EPIMERASE (EC 5.1.3.2) (GALACTOWALDENASE) (UDP-
 DE GALACTOSE 4-EPIMERASE).
 GN GALE.
 OS RATTUS NORVEGICUS (RAT).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SKELETAL MUSCLE;
 RX MEDLINE; 90384840.
 RA ZESCHNICK M., WILCKEN-BERGEMANN B., STARZINSKI-POWITZ A.;
 RL NUCLEIC ACIDS RES. 18:5289-5289(1990).
 CC -1- FUNCTION: CATALYZES TWO DISTINCT BUT ANALOGOUS REACTIONS: THE

Search completed: Thu Jul 30 09:08:29 1998
Job time : 10 secs.

MPSRLH
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd
MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Jul 30 09:07:26 1998; MasPar time 4.27 Seconds
282.006 Million cell updates/sec
Mular output not generated.

Title: >US-08-938-548A-3
Description: (1-33) from US08938548A.pap
Perfect Score: 256
Sequence: 1 QPLPCCRQKTCRLYELLHGAGNHAAGILTL 33
Scoring table: PAM 150
Gap 11
Searched: 120441 seqs, 36531193 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: pir56
1:pir1 2:pir2 3:pir3 4:pir4 5:nrl3d
Statistics: Mean 30.310; Variance 51.743; scale 0.586

pred, No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

				SUMMARIES	
Result No.	Score	Query Match	Length DB ID	Description	Pred. No.
1	72	28.1	147 2 S24303	hypothetical protein	9.01e-01
2	70	27.3	870 2 A41130	dystrophin homolog -	1.84e+00
3	69	27.0	93 2 S60079	oct2 protein isoform	2.63e+00
4	69	27.0	143 2 A37199	sperm mitochondrial c	2.63e+00
5	68	26.6	641 2 JN0800	tetracycline-minocycl	3.73e+00
6	67	26.2	299 1 AHRB	ig alpha chain C regi	5.28e+00
7	67	26.2	338 2 S09276	UDPGlucose 4-epimeras	7.44e+00
8	66	25.8	347 2 S11223	gene D3 protein - mou	7.44e+00
9	66	25.8	425 2 I56329	transcriptional regul	7.44e+00
10	66	25.8	479 2 B69764	hypothetical protein	1.05e+01
11	65	25.4	155 2 C64314	keratin KAP5.5 - shee	1.05e+01
12	65	25.4	197 2 I46413	beta-fructofuranosida	1.05e+01
13	65	25.4	599 2 S33920	Omega-aga-ivb (nmr, m	1.46e+01
14	64	25.0	35 5 LOMA	Omega-aga-ivb (nmr, 2	1.46e+01
15	64	25.0	48 5 IAGG	Omega-aga-ivb (nmr, 2	1.46e+01
16	64	25.0	48 5 IAGG	Omega-aga-ivb (nmr, 2	1.46e+01
17	64	25.0	48 1 A44664	Omega-aga-ivb (nmr, 2	1.46e+01
18	64	25.0	76 2 B54252	Omega-aga-ivb (nmr, 2	1.46e+01
19	64	25.0	76 2 A42335	Omega-aga-ivb (nmr, 2	1.46e+01
20	64	25.0	76 2 A54252	cell division protein	1.46e+01
21	64	25.0	103 2 B64052	pigment deposition co	1.46e+01
22	64	25.0	131 2 A46298	agouti protein precur	1.46e+01
23	64	25.0	132 2 I37143		

24	64	25.0	369	2	B64835	hypothetical protein	1.46e+01
25	64	25.0	493	2	S49175	cysteine proteinase (1.46e+01
26	63	24.6	236	2	H09619	hypothetical protein	2.04e+01
27	63	24.6	360	2	S09271	ig alpha chain C regi	2.04e+01
28	63	24.6	396	2	B57479	amino acid transport	2.04e+01
29	63	24.6	477	2	A29978	angiotensin precursor	2.04e+01
30	63	24.6	639	2	S13142	tetracycline resistan	2.04e+01
31	63	24.6	639	2	A60633	tetracycline resistan	2.04e+01
32	63	24.6	639	2	A56779	tetracycline resistan	2.04e+01
33	62	24.2	157	2	JC5551	hypothetical 17.9K pr	2.84e+01
34	62	24.2	612	2	JH0799	laminin-related prote	2.84e+01
35	62	24.2	722	2	B61231	myosin heavy chain, n	2.84e+01
36	62	24.2	1464	2	S47555	N-methyl-D-aspartate	2.84e+01
37	62	24.2	1464	2	S29159	glutamate receptor, N	2.84e+01
38	62	24.2	1464	2	A43274	N-methyl D-aspartate	2.84e+01
39	62	24.2	1466	2	A36426	SPA2 protein - yeast	2.84e+01
40	61	23.8	76	2	D54252	omega-agatoxin III, 8	3.93e+01
41	61	23.8	160	1	WZBEC7	gene 34 protein - equ	3.93e+01
42	61	23.8	270	2	A42887	urease-associated pro	3.93e+01
43	61	23.8	270	2	S32937	ureb protein - Kiebsi	3.93e+01
44	61	23.8	1169	2	S64859	DNA repair protein RA	3.93e+01
45	61	23.8	2007	1	B43402	myosin heavy chain-B,	3.93e+01

ALIGNMENTS

RESULT 1
ENTRY S24303 #type complete
TITLE hypothetical protein H19-3' - mouse
ORGANISM formal_name Mus musculus #common_name house mouse
DATE 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 10-Sep-1997
ACCESSIONS S24303
REFERENCE S24302
#authors Poirier, F.; Chan, C.T.J.; Timmons, P.M.; Robertson, E.J.; Evans, M.J.; Rigby, P.W.J.
#journal Development (1991) 113:1105-1114
#title The murine H19 gene is activated during embryonic stem cell differentiation in vitro and at the time of implantation in the developing embryo.
#accession S24303
#status preliminary
#molecule_type mRNA
#residues 1-147 #label POI
#cross-references EMBL:X58196; NID:g51131; PID:g51133
SUMMARY #length 147 #molecular-weight 16309 #checksum 6958

Query Match 28.1%; Score 72; DB 2; Length 147;
Best Local Similarity 32.1%; Pred. No. 9.01e-01;
Matches 9; Conservative 9; Mismatches 10; Indels 0; Gaps 0;
Db 66 CCQNHYTTCRLNLLQGEAERTDGVNLL 93
||| : : ||| : : : : :
QY 6 CCRQKTCRLYELLHGAGNHAAGILTL 33
RESULT 2
ENTRY A41130 #type fragment
TITLE dystrophin homolog - Pacific electric ray (fragment)
ALTERNATE_NAMES 300k subsynaptic protein
ORGANISM #formal_name Torpedo californica #common_name Pacific electric ray
DATE 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 23-Jan-1998
ACCESSIONS A41130
REFERENCE A41130
#authors Yeadon, J.E.; Lin, H.; Dyer, S.M.; Burden, S.J.
#journal J. Cell Biol. (1991) 115:1069-1076
#title Dystrophin is a component of the subsynaptic membrane.
#cross-references MUID:92064638
#accession A41130
#molecule_type mRNA
#residues 1-870 #label YEA

stabilized by interchain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger oligomers.

CLASSIFICATION #superfamily immunoglobulin C region; immunoglobulin homology
KEYWORDS duplication; glycoprotein; heterotetramer; immunoglobulin; plasma

FEATURE #domain immunoglobulin homology #label IGG1\
86-152 #domain immunoglobulin homology #label IGG2\
189-261 #binding_site carbohydrate (Asn) (covalent) #status
38,286 predicted

SUMMARY #length 299 #checksum 2361

Query Match 26.2%; Score 67; DB 1; Length 299;
Best Local Similarity 50.0%; Pred. No. 5.28e+00;

Matches 7; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Db 44 PFPDCCPANSCTC 57

QY 2 PFPDCCQKTC-SC 14

RESULT 7

ENTRY #type fragment

TITLE Ig alpha chain C region - rabbit (fragment)

ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic rabbit

DATE 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change

ACCESSIONS S09276

REFERENCE S09264

#authors Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.

#journal EMBO J. (1989) 8:4041-4047

#title The IgA heavy-chain gene family in rabbit: cloning and

sequence analysis of 13 C-alpha genes.

#cross-references MUID:90076124

#accession S09276

#status not compared with conceptual translation

#molecule_type DNA

#residues 1-338 #label BUR

CLASSIFICATION #superfamily immunoglobulin C region; immunoglobulin homology

KEYWORDS immunoglobulin

FEATURE #domain immunoglobulin homology #label IGG2

228-300 #length 338 #checksum 2169

Query Match 26.2%; Score 67; DB 2; Length 338;

Best Local Similarity 50.0%; Pred. No. 5.28e+00;

Matches 7; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Db 83 PFPDCCPANSCTC 96

QY 2 PFPDCCQKTC-SC 14

RESULT 8

ENTRY #type complete

TITLE UDPglucose 4-epimerase (EC 5.1.3.2) - rat

ALTERNATE_NAMES UDPgalactose 4-epimerase

ORGANISM #formal_name Rattus norvegicus #common_name Norway rat

DATE 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change

08-Sep-1997

ACCESSIONS S11223

REFERENCE S11223

#authors Zeschnick, M.; von Wilcken-Bergmann, B.; Starzinski-Powitz, A.

#journal Nucleic Acids Res. (1990) 18:5289

#title cDNA from rat cells with constitutive galactose-epimerase

activity in E. coli.

#cross-references MUID:90384840

#accession S11223

#status preliminary

#molecule_type mRNA

#residues 1-347 #label ZES
#cross-references EMBL:X53949; NID:g57791; PID:g57792

GENETICS

#gene gale

CLASSIFICATION #superfamily Escherichia coli UDPglucose 4-epimerase;

KEYWORDS UDPglucose 4-epimerase homology

FEATURE galactose metabolism; isomerase

SUMMARY #domain UDPglucose 4-epimerase homology #label UDP
#length 347 #molecular-weight 38225 #checksum 6947

Query Match 25.8%; Score 66; DB 2; Length 347;

Best Local Similarity 53.8%; Pred. No. 7.44e+00;

Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 261 CGCRIYNLGTGTG 273

QY 12 CSCRLYELLHGAG 24

RESULT 9

ENTRY I56329 #type complete

TITLE gene D3 protein - mouse

ORGANISM #formal_name Mus sp. #common_name mouse

DATE 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change

ACCESSIONS I56329

REFERENCE I56329

#authors Tannenbaum, C.S.; Major, J.; Ohmori, Y.; Hamilton, T.A.

#journal J. Leukoc. Biol. (1993) 53:563-568

#title A lipopolysaccharide-inducible macrophage gene (D3) is a new member of an interferon-inducible gene cluster and is selectively expressed in mononuclear phagocytes.

#cross-references MUID:93274206

#accession I56329

#status preliminary; translated from GB/EMBL/DBDJ

#molecule_type mRNA

#residues 1-425 #label RES

#cross-references GB:S62227; NID:g385702; PID:g385703

GENETICS

#gene D3

SUMMARY #length 425 #molecular-weight 47046 #checksum 8487

Query Match 25.8%; Score 66; DB 2; Length 425;

Best Local Similarity 31.0%; Pred. No. 7.44e+00;

Matches 9; Conservative 10; Mismatches 9; Indels 1; Gaps 1;

Db 178 PCCCEPTAMCQ-SPILHSSSSASSNNILS 205

QY 4 PDCCRKTCRDLHLLHGAGNHAAGILT 32

RESULT 10

ENTRY B69764 #type complete

TITLE transcriptional regulator (GntR family) / homolog ycnF -

ORGANISM Bacillus subtilis

DATE #formal_name Bacillus subtilis

05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change

ACCESSIONS B69764

REFERENCE A69580

#authors

Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;

Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.;

Bolotin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans,

A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;

Bruschl, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;

Choi, S.K.; Codani, J.J.; Connerton, I.F.; Cummings, N.J.;

Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoeft, A.;

Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.;

Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita,

M.; Fujita, Y.; Funa, S.; Galizzi, A.; Galleron, N.; Ghim,

S.Y.; Glaser, P.; Goffeau, A.; Golligly, E.J.; Grandi, G.;

Guiseppi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,

1-23 #domain signal sequence #status experimental #label sig\
24-589 #product beta-fructofuranosidase #status experimental
#label MAT\
36,42,170,188,211,
254,259,318,322,
388,463,518,527
#binding_site carbohydrate (Asn) (covalent) #status
predicted\
70,92 #binding_site phosphate (Thr) (covalent) #status
predicted\
458,475,490 #binding_site phosphate (Ser) (covalent) #status
predicted
SUMMARY #length 589 #molecular-weight 63650 #checksum 1032
Query Match 25.4%; Score 65; DB 2; Length 589;
Best Local Similarity 42.1%; Pred. No. 1.05e+01;
Matches 8; Conservative 6; Mismatches 4; Indels 1; Gaps 1;
Db 548 RLFDVING-GEQAIETL 565
||::||| ||::||
QY 15 RLVELLHGAGNHAAGILYL 33

RESULT 14
ENTRY 1OMB #type complete
TITLE Omega-aga-ivb (nmr, minimized average structure) -
#funnel-weaving spider (Agelenopsis aperta)
ORGANISM #formal_name Agelenopsis aperta
REFERENCE A51322
#authors Yu, H.; Rosen, M.K.; Schreiber, S.L.
#submission submitted to the Brookhaven Protein Data Bank, September 1993
#cross-references PDB:1OMB
COMMENT Resolution: not applicable
COMMENT Determination: NMR
FEATURE
7-9,31-33,23-24 #region beta sheet\
10-13 #region turn (type II)\
1-17 #disulfide_bonds\
9-22 #disulfide_bonds\
16-33 #disulfide_bonds\
24-31 #disulfide_bonds\
SUMMARY #length 35 #molecular-weight 3804 #checksum 7419

Query Match 25.0%; Score 64; DB 5; Length 35;
Best Local Similarity 45.5%; Pred. No. 1.46e+01;
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 16 CCRGRPCRCSM 26
|||::| | |
QY 6 CCRQKTCSCRL 16

RESULT 15
ENTRY 1OMA #type complete
TITLE Omega-aga-ivb (nmr, 21 structures) - funnel-weaving spider
ORGANISM #formal_name Agelenopsis aperta
REFERENCE A51322
#authors Yu, H.; Rosen, M.K.; Schreiber, S.L.
#submission submitted to the Brookhaven Protein Data Bank, September 1993
#cross-references PDB:1OMA
COMMENT Resolution: not applicable
COMMENT Determination: NMR
FEATURE
10-12,34-36,26-27 #region beta sheet\
13-16 #region turn (type II)\
4-20 #disulfide_bonds\
12-25 #disulfide_bonds\
19-36 #disulfide_bonds\
27-34 #disulfide_bonds\
SUMMARY #length 48 #molecular-weight 5281 #checksum 7904

Query Match 25.0%; Score 64; DB 5; Length 48;
Best Local Similarity 45.5%; Pred. No. 1.46e+01;

Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
Db 19 CCRGRPCRCSM 29
|||::| | |
QY 6 CCRQKTCSCRL 16
Search completed: Thu Jul 30 09:08:01 1998
Job time : 35 secs.

CC FILING DATE: 03-SEP-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/907,138
 CC FILING DATE: 30-JUN-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/863,703
 CC FILING DATE: 03-APRIL-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: U.K. 91 07566.3
 CC FILING DATE: 10-APRIL-1991
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Hanson, Norman D.
 CC REGISTRATION NUMBER: 30,946
 CC REFERENCE/DOCKET NUMBER: LUD 5250.5
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (212) 688-9200
 CC TELEFAX: (212) 838-3884
 CC INFORMATION FOR SEQ ID NO: 40:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 39
 CC TYPE: amino acid
 CC STRANDEDNESS:
 CC TOPOLOGY: linear
 CC SEQUENCE 39 AA; 3760 MW; 6521 CN;
 CC
 CC Query Match 25.8%; Score 66; DB 2; Length 39;
 CC Best Local Similarity 70.0%; Pred. No. 2.23e+01;
 CC Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
 CC
 CC Db 8 CCR-TTCACR 16
 CC III III
 CC QY 6 CCRQKTCSCR 15
 CC
 CC RESULT 4
 CC ID PCT-US94-05083C-40 STANDARD; PRT; 39 AA.
 CC AC XXXXXX
 CC XX
 CC DT
 CC XX
 CC DE Sequence 40, Application PC/TUS9405083C
 CC XX
 CC Sequence 40, Application PC/TUS9405083C
 CC GENERAL INFORMATION:
 CC APPLICANT: Robert Sklar, Mark Marchionni,
 CC APPLICANT: David I. Gwynne
 CC TITLE OF INVENTION: METHODS FOR ALTERING
 CC TITLE OF INVENTION: MUSCLE CONDITION
 CC NUMBER OF SEQUENCES: 185
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Fish & Richardson
 CC STREET: 225 Franklin Street
 CC CITY: Boston
 CC STATE: Massachusetts
 CC ZIP: 02110-2804
 CC COMPUTER: IBM
 CC OPERATING SYSTEM: PC-DOS
 CC SOFTWARE: Wordperfect
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: PCT/US94/05083C
 CC FILING DATE: 06-MAY-94
 CC CLASSIFICATION:
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/209,204
 CC FILING DATE: 08-MAR-94
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/059,022
 CC FILING DATE: 06-May-93
 CC ATTORNEY/AGENT INFORMATION:

CC NAME: Clark, Paul T.
 CC REGISTRATION NUMBER: 30,162
 CC REFERENCE/DOCKET NUMBER: 04585/028W01
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 542-5070
 CC TELEFAX: (617) 542-8906
 CC TELEX: 200154
 CC INFORMATION FOR SEQ ID NO: 40:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 39
 CC TYPE: amino acid
 CC STRANDEDNESS:
 CC TOPOLOGY: linear
 CC SEQUENCE 39 AA; 3760 MW; 6521 CN;
 CC
 CC Query Match 25.8%; Score 66; DB 2; Length 39;
 CC Best Local Similarity 70.0%; Pred. No. 2.23e+01;
 CC Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
 CC
 CC Db 8 CCR-TTCACR 16
 CC III III
 CC QY 6 CCRQKTCSCR 15
 CC
 CC RESULT 5
 CC ID US-08-469-569-40 STANDARD; PRT; 39 AA.
 CC AC XXXXXX
 CC XX
 CC DT
 CC XX
 CC DE Sequence 40, Application US/08469569
 CC XX
 CC Sequence 40, Application US/08469569
 CC Patent No. 5606032
 CC GENERAL INFORMATION:
 CC APPLICANT: Goodearl, Andrew; Stroobant, Paul;
 CC APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;
 CC APPLICANT: Chen, Maio Su; Hiles, Ian
 CC TITLE OF INVENTION: Gli1 Mitogenic Factors, Their
 CC TITLE OF INVENTION: Preparation and Use
 CC NUMBER OF SEQUENCES: 184
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Felfe & Lynch
 CC STREET: 805 Third Avenue
 CC CITY: New York City
 CC STATE: New York
 CC COUNTRY: USA
 CC ZIP: 10022
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
 CC COMPUTER: IBM
 CC OPERATING SYSTEM: PC-DOS
 CC SOFTWARE: Wordperfect
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/469,569
 CC FILING DATE: 06-JUN-1995
 CC CLASSIFICATION: 530
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/036,555
 CC FILING DATE: 24-MAR-1993
 CC APPLICATION NUMBER: 07/965,173
 CC FILING DATE: 23-OCT-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/940,389
 CC FILING DATE: 03-SEP-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/907,138
 CC FILING DATE: 30-JUN-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/863,703
 CC FILING DATE: 03-APRIL-1992
 CC PRIOR APPLICATION DATA:

CC REFERENCE/DOCKET NUMBER: MA48DD2.C1
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (904) 375-8100
 CC TELEFAX: (904) 372-3800
 CC INFORMATION FOR SEQ ID NO: 6:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 1167 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 1167 AA; 131657 MW; 7095219 CN;
 SQ
 Query Match 25.0%; Score 64; DB 1; Length 1167;
 Best Local Similarity 45.5%; Pred. No. 3.29e+01;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Db 873 PNCCKPAACOC 883
 I:|:|:|:|:|
 4 PDCCRQKTCSC 14
 CC
 RESULT 8
 ID US-08-620-717A-9 STANDARD; PRT; 1168 AA.
 CC
 AC xxxxxx
 CC
 XX
 CC
 DT
 CC
 XX
 CC
 DE
 CC
 XX
 CC
 Sequence 9, Application US/08620717A
 Sequence 9, Application US/08620717A
 Patent No. 5670365
 GENERAL INFORMATION:
 CC APPLICANT: Feitelson, Jerald S.
 CC TITLE OF INVENTION: Identification of, and Uses For, Nematicidal
 CC TITLE OF INVENTION: Bacillus thuringiensis Genes, Toxins, and Isolates
 CC NUMBER OF SEQUENCES: 9
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Saliwanchik & Saliwanchik
 CC STREET: 2421 N.W. 41st Street, Suite A-1
 CC CITY: Gainesville
 CC STATE: Florida
 CC COUNTRY: USA
 CC ZIP: 32606
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/620,717A
 CC FILING DATE:
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/540,104
 CC FILING DATE: 06-OCT-1995
 CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Saliwanchik, David R.
 CC REGISTRATION NUMBER: 31,794
 CC REFERENCE/DOCKET NUMBER: MA94.C1
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (352) 375-8100
 CC TELEFAX: (352) 372-5800
 CC INFORMATION FOR SEQ ID NO: 9:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 1168 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC ORIGINAL SOURCE:

CC INDIVIDUAL ISOLATE: 167P
 SQ SEQUENCE 1168 AA; 131542 MW; 7175894 CN;
 Query Match 25.0%; Score 64; DB 1; Length 1168;
 Best Local Similarity 45.5%; Pred. No. 3.29e+01;
 Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Db 873 PNCCKPAACOC 883
 I:|:|:|:|:|
 4 PDCCRQKTCSC 14
 CC
 RESULT 9
 ID US-08-026-138E-1 STANDARD; PRT; 1464 AA.
 CC
 AC xxxxxx
 CC
 XX
 CC
 DT
 CC
 XX
 CC
 DE
 CC
 XX
 CC
 Sequence 1, Application US/08026138E
 Sequence 1, Application US/08026138E
 Patent No. 5502166
 GENERAL INFORMATION:
 CC APPLICANT: Masayoshi MISHINA
 CC TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING THE SAME
 CC NUMBER OF SEQUENCES: 19
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Nishiohata Residence 1-107
 CC STREET: 5214, Nishiohata-machi
 CC CITY: Niigata-shi
 CC STATE: Niigata-ken
 CC COUNTRY: JAPAN
 CC ZIP: 951
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
 CC COMPUTER: IBM Compatible
 CC OPERATING SYSTEM: MS-DOS V.5
 CC SOFTWARE: Word Perfect 5.1
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/026,138E
 CC FILING DATE: 26-FEB-1993
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: JP 39563/1992
 CC FILING DATE: 26-FEB-1992
 CC APPLICATION NUMBER: JP 173155/1992
 CC FILING DATE: 30-JUN-1992
 CC APPLICATION NUMBER: JP 215017/1992
 CC FILING DATE: 12-AUG-1992
 CC APPLICATION NUMBER: JP 303878/1992
 CC FILING DATE: 13-NOV-1992
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Hamburg, C. Bruce
 CC REGISTRATION NUMBER: 22,389
 CC REFERENCE/DOCKET NUMBER: F-4551
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (212) 986-2340
 CC TELEFAX: (212) 953-7733
 CC INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 1464 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single strand
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC ORIGINAL SOURCE:
 CC ORGANISM: mouse
 CC TISSUE TYPE: cerebellum
 CC PUBLICATION INFORMATION:
 CC AUTHORS: Masayoshi MISHINA
 CC TITLE: NOVEL PROTEINS AND GENES CODING THE SAME
 CC RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 to 1464
 CC SEQUENCE 1464 AA; 165489 MW; 11224000 CN;
 SQ


```

RESULT 12
ID US-08-477-383-23 STANDARD; PRT; 42 AA.
XX
AC
XX
DT
XX
DE
XX
Sequence 23, Application US/08477383
Sequence 23, Application US/08477383
Patent No. 5589340
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hilliard, David R.
APPLICANT: Macintosh, J. Michael
APPLICANT: Santos, Ameurfino S.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,383
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/137,800
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/084,848
FILING DATE: 29-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Innen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107673
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-8300
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Conus ochroleucus
SQ SEQUENCE 42 AA; 4618 MW; 9505 CN;

Query Match 23.8%; Score 61; DB 1; Length 42;
Best Local Similarity 47.4%; Pred. No. 5.85e+01;
Matches 9; Conservative 3; Mismatches 5; Indels 2; Gaps 2;

Db 14 PVPSCCQVSSCW-NLYGL 31
|:| |:| |:| |:|
QY 2 PLP-DCCRQKTCSCRLYL 19

RESULT 13
ID US-08-137-800-23 STANDARD; PRT; 42 AA.
XX
AC
XX
DT
XX
DE
XX
Sequence 23, Application US/08477383
Sequence 23, Application US/08477383
Patent No. 5589340
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hilliard, David R.
APPLICANT: Macintosh, J. Michael
APPLICANT: Santos, Ameurfino S.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,383
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/137,800
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/084,848
FILING DATE: 29-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Innen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107673
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-8300
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Conus ochroleucus
SQ SEQUENCE 42 AA; 4618 MW; 9505 CN;

Query Match 23.8%; Score 61; DB 1; Length 42;
Best Local Similarity 47.4%; Pred. No. 5.85e+01;
Matches 9; Conservative 3; Mismatches 5; Indels 2; Gaps 2;

Db 14 PVPSCCQVSSCW-NLYGL 31
|:| |:| |:| |:|
QY 2 PLP-DCCRQKTCSCRLYL 19

```

```

XX
AC
XX
DT
XX
DE
XX
Sequence 23, Application US/08137800
Sequence 23, Application US/08137800
Patent No. 5514774
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hilliard, David R.
APPLICANT: Macintosh, J. Michael
APPLICANT: Santos, Ameurfino D.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue N.W., Suite 1000
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137,800
FILING DATE: 19-OCT-1993
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Innen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-104763
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Conus ochroleucus
SQ SEQUENCE 42 AA; 4618 MW; 9505 CN;

Query Match 23.8%; Score 61; DB 1; Length 42;
Best Local Similarity 47.4%; Pred. No. 5.85e+01;
Matches 9; Conservative 3; Mismatches 5; Indels 2; Gaps 2;

Db 14 PVPSCCQVSSCW-NLYGL 31
|:| |:| |:| |:|
QY 2 PLP-DCCRQKTCSCRLYL 19

RESULT 14
ID US-08-233-788A-51 STANDARD; PRT; 59 AA.
XX
AC
XX
DT
XX
Sequence 51, Application US/08233788A
Sequence 51, Application US/08233788A
Patent No. 5635617
GENERAL INFORMATION:
APPLICANT: Doran, James L.

```

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
 on: Thu Jul 30 09:06:53 1998; MasPar time 3.32 Seconds
 152.287 Million cell updates/sec
 ular output not generated.

```
>US-08-938-548A-3
Description: (1-33) from US08938548A.pep
Perfect Score: 256
Sequence: 1 QPLPCCCRKQKTCSCRLYELLHGAGNHAAGILT 33
```

Scoring table: PAM 150
Gap 11

Searched: 124785 seqs, 15338987 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

a-genes31-2

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27

Statistics: Mean 22.991; Variance 84.736; scale 0.271

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	77	30.1	373	23	W12414	Porcine complement in	6.26e+00
2	75	29.3	373	18	R98461	Murine ICE-ced-3 homo	9.46e+00
3	75	29.3	373	13	R66767	Murine interleukin-1	9.46e+00
4	69	27.0	329	26	W9877	Lysophosphatidic acid	3.20e+01
5	66	25.8	102	1	X95679	Xenopus Vg1 protein f	5.80e+01
6	66	25.8	348	20	W01619	Human uridine diphosph	5.80e+01
7	65	25.4	78	5	P20020	Sequence of a foot an	7.06e+01
8	64	25.0	45	23	W10106	Human agouti signall	8.59e+01
9	64	25.0	45	23	W10105	Murine agouti signall	8.59e+01
10	64	25.0	48	9	R45611	AG1 toxin.	8.59e+01
11	64	25.0	48	11	R60293	Calcium channel inhib	8.59e+01
12	64	25.0	48	8	R44209	A. aperta venom fract	8.59e+01
13	64	25.0	130	23	W10102	Human agouti signall	8.59e+01
14	64	25.0	131	23	W10101	Murine agouti signall	8.59e+01
15	64	25.0	1167	27	W31504	Nematode toxin 167p p	8.59e+01
16	64	25.0	1167	20	W0653	Bacillus thuringiens	8.59e+01
17	64	25.0	1168	23	W63326	Nematocidal toxin 167	8.59e+01
18	62	24.2	34	10	R55058	Tarantula spider veno	1.27e+02
19	62	24.2	1464	10	R35529	Human NMDA R2A recept	1.27e+02

ALIGNMENTS

RESULT	
ID	W12414 standard; Protein; 363 AA.
AC	W12414;
DE	24-SEP-1997 (first entry)
DT	Porcine complement inhibitor.
DE	porcine; pig; complement; inhibitor; organ transplantation;
KW	analysis; promoter.
KW	Sus scrofa.
OS	Sus scrofa.
PN	WC9700951-Al.
PD	09-JAN-1997.
PF	19-JUN-1996; J01704.
PR	19-JUN-1995; JP-178254.
PA	(NINE-) NIPPON MEAT PACKERS INC.
PI	(NIHA-) NIPPON HAM KK.
PI	Murakami H, Shigehisa T, Toyomura K;
DR	WPI: 97-087378/08.
DR	N-PSDB: T61098.
PT	DNA encoding porcine complement inhibitor - useful in porcine organ
PT	transplant to humans
PS	Clam 3; Page 12-14; 20pp; Japanese.
CC	This protein is a porcine complement inhibitor encoded by pMCPcDNA
CC	(T61098). The DNA is useful for large scale production of
CC	recombinant porcine complement inhibitor, which is useful for
CC	recaine organ transplantation into humans. The DNA clone pMCPcDNA is
CC	also useful in the analysis of the promoter region of porcine complement
CC	inhibitor.
SQ	Sequence 363 AA;
Query Match	30.1%; Score 77; DB 23; Length 363;
Best Local Similarity	39.3%; Pred. No. 6.26e+00;
Matches	11; Conservative 7; Mismatches 9; Indels 1; Gaps 1;

Human N-methyl-D-aspa	1.27e+02
Human excitatory amin	1.27e+02
Glutamic acid recepto	1.27e+02
Rat NMDA receptor sub	1.27e+02
A-ligand conotoxin p	1.53e+02
Predatory cone snail	1.53e+02
Salmonella enteritidis	1.53e+02
Human membrane antige	1.53e+02
Tctc sequence.	1.53e+02
Human bg protein asso	1.53e+02
Human bg protein asso	1.53e+02
Cyn d allergen B1.	1.86e+02
Cherry polyphenol oxi	1.86e+02
E. coli DNA polymeras	1.86e+02
APP-HCV-2E fusion pro	1.86e+02
Spider venom peptide	2.24e+02
N-terminal sequence o	2.24e+02
Funnel-web spider ven	2.24e+02
Connexin-32.	2.24e+02
Sequence of viper ven	2.71e+02
Sequence of viper ven	2.71e+02
Ovine FSH beta subuni	2.71e+02
Ikaros protein.	2.71e+02
Mouse 22B/30B (candid	2.71e+02
Murine Lysyl1 long iso	2.71e+02

[illegible]

PR 18-AUG-1981; GB-023130.

ID W10105 standard; protein;
AC W10105;
DT 17-SEP-1997 (first entry)

PS Example; Page 17; 28pp; English.
 CC The sequence is that of a polypeptide present in fraction K of the
 CC venom of *Agelenopsis aperta*, it blocks calcium channels in cells
 CC of both mammals and invertebrates, partic. those affecting neuronal
 CC and muscle cells. It may be used in the treatment of angina,
 CC hypertension, cardiomyopathies, supraventricular arrhythmia,
 CC oesophageal achalasia, premature labour, and Raynaud's disease.
 CC It may also be of use in the study of cell physiology and in the
 CC control of invertebrate pests. It may be produced synthetically.
 SQ Sequence 48 AA;

Query Match 25.0%; Score 64; DB 8; Length 48;
 Best Local Similarity 45.5%; Pred. No. 8.59e+01;
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 19 corgprcscsm 29
 QY 6 CCRQKTCSCRL 16
 ||| : ||| :

ULT 13
 AC W10102 standard; protein; 130 AA.

DE Human agouti signalling protein.
 DT 17-SEP-1997 (first entry)
 KW Agouti signalling protein; ASP; depigmenting activity; cosmetic;
 KW hyperpigmentary condition; melasma photoageing spots; solar keratosis;
 KW post-inflammatory hyperpigmentation; wound healing; eumelanogenesis;
 KW vitiligo; leucoderma; albinism; hair greying.

OS Homo sapiens.

PH Key Location/Qualifiers

FT peptide 1..21 /note= "Signal sequence"

FT region 32

FT region 86 /note= "Start of basic region"

FT region 86 /note= "Start of Cysteine-rich motif"

FT WO9700892-A2.

PD 09-JAN-1997.

PF 21-JUN-1996; U10695.

PR 23-JUN-1995; US-000436.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Hearing VJ;

DR WPI; 97-087323/08.

PT New agouti signal protein peptide(s) and nucleic acids - used for
 PT altering melanin prodn., for treating e.g. melasma photo-ageing

PT spots, solar keratosis or vitiligo

PS Claim 6; Page 8-9; 67pp; English.

CC The sequences given in W10101-29 are biologically active peptides
 CC and fragments of the agouti signalling protein (ASP) which have
 CC depigmenting activity. These peptides are useful for cosmetic purposes
 CC and for clinical application in the prevention or treatment of various
 CC hyperpigmentary conditions and diseases such as melasma photoageing

CC spots, solar keratosis, and post-inflammatory hyperpigmentation such as
 CC occurs at sites of wound healing. They can also be used to provide
 CC enhanced eumelanogenesis for treating eg. vitiligo, leucoderma, some
 CC forms of albinism and hair greying.

SQ Sequence 130 AA;

Query Match 25.0%; Score 64; DB 23; Length 130;
 Best Local Similarity 41.2%; Pred. No. 8.59e+01;
 Matches 7; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

Db 101 papaccdpascqcrff 117

QY 2 PLPDCCRQ-KTCSCLRY 17
 ||| : ||| :

RESULT 14

ID W10101 standard; protein; 131 AA.

AC W10101;

DT 17-SEP-1997 (first entry)

DE Murine agouti signalling protein.

KW Agouti signalling protein; ASP; depigmenting activity; cosmetic;
 KW hyperpigmentary condition; melasma photoageing spots; solar keratosis;
 KW post-inflammatory hyperpigmentation; wound healing; eumelanogenesis;
 KW vitiligo; leucoderma; albinism; hair greying.

OS Mus musculus.

PH Key Location/Qualifiers

FT peptide 1..21 /note= "Signal sequence"

FT region 32 /note= "Start of basic region"

FT region 87

FT /note= "Start of Cysteine-rich motif"

FT WO9700892-A2.

PD 09-JAN-1997.

PF 21-JUN-1996; U10695.

PR 23-JUN-1995; US-000436.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Hearing VJ;

DR WPI; 97-087323/08.

PT New agouti signal protein peptide(s) and nucleic acids - used for
 PT altering melanin prodn., for treating e.g. melasma photo-ageing

PT spots, solar keratosis or vitiligo

PS Claim 5; Page 8-9; 67pp; English.

CC The sequences given in W10101-29 are biologically active peptides
 CC and fragments of the agouti signalling protein (ASP) which have
 CC depigmenting activity. These peptides are useful for cosmetic purposes
 CC and for clinical application in the prevention or treatment of various
 CC hyperpigmentary conditions and diseases such as melasma photoageing

CC spots, solar keratosis, and post-inflammatory hyperpigmentation such as
 CC occurs at sites of wound healing. They can also be used to provide
 CC enhanced eumelanogenesis for treating eg. vitiligo, leucoderma, some
 CC forms of albinism and hair greying.

SQ Sequence 131 AA;

Query Match 25.0%; Score 64; DB 23; Length 131;
 Best Local Similarity 41.2%; Pred. No. 8.59e+01;
 Matches 7; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

Db 102 papaccdpascqcrff 118

QY 2 PLPDCCRQ-KTCSCLRY 17
 ||| : ||| :

RESULT 15

ID W31504 standard; Protein; 1167 AA.

AC W31504;

DT 07-APR-1998 (first entry)

DE Nematode toxin 167p protein.

KW PCR primer; amplify; nematode toxic protein; *Bacillus thuringiensis*;

KW delta-endotoxin gene; nematode pest control; *Panagrellus redivivus*;

KW 167P protein.

OS *Bacillus thuringiensis*.

PN WO9734926-A2.

PD 25-SEP-1997.

PF 21-MAR-1997; U04755.

PR 21-MAR-1996; US-590554.

PA (MYCO) MYCOGEN CORP.

PI Fu J, Narva KE, Payne J;

DR WPI; 97-480163/44.

DR N-PSDB; T89185.

PT *Bacillus thuringiensis* toxin gene - useful in recombinant hosts,

PT particularly plants for the control of nematodes

PS Claim 4; Page 35-39; 44pp; English.

CC This sequence represents the protein encoded by a polynucleotide of the

CC invention. The polynucleotide of the invention is a sequence from a

CC *Bacillus thuringiensis* (Bt) isolate selected from PS80J1, PS158D5,

CC PS167P, PS169E, PS177F1, PS177G, PS204E4 and PS204G6, that encodes a

CC toxin active against nematodes. This sequence represents the 167P

CC protein, and is a delta-endotoxin protein. The polynucleotides and toxins

CC can be used for the control of nematode pests such as *Panagrellus*

CC redivivus. 1167 AA;

SQ

WIREH (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:04:06 1998; MasPar time 9.31 Seconds
592.652 Million cell updates/sec

Linear output not generated.

Title: >US-08-938-548A-2
Description: (1-131) from US08938548A.pep
Perfect Score: 931
Sequence: 1 MNLPTKVSNAVTLTLL.....GRRCSAPAAASVAPGGSGI 131

Scoring table:
PAM 150
Gap 11

Searched: 140542 seqs, 42109429 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrembl5
1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mnc 6:sp_organelle 7:sp_phase 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

Statistics: Mean 39.725; Variance 90.808; scale 0.437

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Description	ID	Pred. No.
1	109	11.7	PROTAGLANDIN H SYNTHA	Q62731	2.09e-02
2	108	11.6	PROTAGLANDIN ENDOPERO	Q15122	2.80e-02
3	107	11.5	CYCLOOXYGENASE 1.	Q63684	3.74e-02
4	107	11.5	CYCLOOXYGENASE ISOFORM	Q63921	3.74e-02
5	105	11.3	FROM BASES 2573751 TO	P76559	6.65e-02
6	103	11.1	HYPOTHETICAL 74.6 KD P	Q23352	1.18e-01
7	101	10.8	BAI 1.	Q14514	2.07e-01
8	96	10.3	NISHED (FRAGMENT).	Q42394	8.25e-01
9	95	10.2	HYPOTHETICAL 23.1 KD P	Q08319	1.08e+00
10	94	10.1	PROTEIN-TYROSINE PHOSP	Q15255	1.42e+00
11	93	10.0	ORF459 PROTEIN.	Q37839	1.86e+00
12	93	10.0	GUANYLATE CYCLASE E.	Q19179	1.86e+00
13	92	9.9	ANDROGEN-BINDING PROTE	Q97518	2.42e+00
14	92	9.9	SEX HORMONE-BINDING GL	Q62588	2.42e+00
15	92	9.9	SORTILIN PRECURSOR.	Q99523	2.42e+00
16	90	9.7	3' ORF.	Q61639	4.11e+00
17	90	9.7	ACYLOXYACYL HYDROLASE.	Q35298	4.11e+00
18	90	9.7	LEUCINE-RICH REPEAT/RE	Q40699	4.11e+00
19	89	9.6	FLT3 LIGAND, T169 FORM	Q61104	5.34e+00
20	89	9.6	PRISTINAMYCIN I SYNTHA	Q07944	5.34e+00

21	88	9.5	480	2	Q92743	NOVEL SERINE PROTEASE.	6.93e+00
22	88	9.5	519	1	Q13420	PHENOLOXIDASE (EC 1.10	6.93e+00
23	88	9.5	1174	4	Q95168	TIGHT JUNCTION PROTEIN	6.93e+00
24	87	9.3	84	10	Q60471	ANION EXCHANGER ISOFOR	8.96e+00
25	87	9.3	551	9	P72405	PCBR.	8.96e+00
26	87	9.3	729	10	Q60470	ANION EXCHANGER 2 A (F	8.96e+00
27	87	9.3	1534	4	Q28298	RIBOSOME RECEPTOR.	1.16e+01
28	86	9.2	142	4	Q28334	INTERLEUKIN-3.	1.16e+01
29	86	9.2	256	9	Q07405	MAY266 (FRAGMENT).	1.16e+01
30	86	9.2	277	9	Q33285	HYPOTHETICAL 30.9 KD P	1.16e+01
31	86	9.2	2195	1	Q02822	VESICLE COAT PROTEIN S	1.16e+01
32	85	9.1	100	8	Q43535	LIM4 PRECURSOR (FRAGME	1.49e+01
33	85	9.1	139	9	Q05844	VRG53 PROTEIN (FRAGEN	1.49e+01
34	85	9.1	143	2	Q15412	CTG4A.	1.49e+01
35	85	9.1	233	2	Q14696	MNA (KIAA0081) FOR OR	1.49e+01
36	85	9.1	331	12	Q91640	LEUCINE ZIPPER WITH BA	1.49e+01
37	85	9.1	363	4	Q02839	PORCINE MEMBRANE COFAC	1.49e+01
38	85	9.1	422	2	Q07111	GLIAL GROWTH FACTOR 2	1.49e+01
39	85	9.1	760	10	Q62178	SEMAPHORIN B.	1.49e+01
40	85	9.1	767	11	Q66627	ORF 24.	1.49e+01
41	84	9.0	285	10	Q35083	1-ACYL-SN-GLYCEROL-3-P	1.92e+01
42	84	9.0	438	9	Q51365	O-POLYMERASE RFC (RFC)	1.92e+01
43	84	9.0	440	10	Q35849	O-POLYMERASE RFC (RFC)	1.92e+01
44	84	9.0	793	2	Q14968	LECITHIN:CHOLESTEROL A	1.92e+01
45	84	9.0	2133	11	Q98203	APOR2DELTA4-7.	1.92e+01
					MC035R.		1.92e+01

ALIGNMENTS

RESULT 1	PRELIMINARY; PRT; 602 AA.
ID Q62731	
AC Q62731	
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)	
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)	
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)	
DE PROTAGLANDIN H SYNTHASE.	
GN PGHS-1.	
OS RATUUS NORVEGICUS (RAT).	
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;	
OC EUTHERIA; RODENTIA.	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN-FISCHER 344; TISSUE-TRACHEA;	
RX MEDLINE; 9516876.	
RA KITZLER J., HILL E., HARDMAN R., REDDY N., PHILPOT R., ELING T.E.;	
RL ARCH. BIOCHEM. BIOPHYS. 316:856-863(1995).	
DR EMBL; U18050; G603052.	
SQ SEQUENCE 602 AA; 69032 MW; 7E3888D7 CRC32;	

Query Match 11.7%; Score 109; DB 10; Length 602;
Best Local Similarity 53.6%; Pred. No. 2.09e-02;
Matches 15; Conservative 6; Mismatches 5; Indels 2; Gaps 1;

RESULT 2	PRELIMINARY; PRT; 599 AA.
ID Q15122	
AC Q15122	
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)	
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)	
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)	
DE PROTAGLANDIN ENDOPEROXIDE SYNTHASE.	
OS HOMO SAPIENS (HUMAN).	
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;	
OC EUTHERIA; PRIMATES.	
RN [1]	
RP SEQUENCE FROM N.A.	
RX MEDLINE; 90088508.	
RA YOKOYAMA C., TANABE T.;	
RL BIOCHEM. BIOPHYS. RES. COMMUN. 165:888-894(1989).	

Db 655 SP 656
QY 102 GA 103

RESULT 7
ID O14514 PRELIMINARY; PRT; 1584 AA.
AC O14514;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE BAI 1.
GN BAI 1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA NISHIMORI H., SHIRATSUCHI T., URANO T., KIMURA Y., KIYONO K.,
TATSUMI K., YOSHIDA S., ONO M., KUWANO M., NAKAMURA Y.,
ONCOGENE 0:0-0(1997).
DR EMBL; AB005297; D1024528;
SQ SEQUENCE 1584 AA; 173531 MW; 235A5C42 CRC32;

Query Match 10.8%; Score 101; DB 2; Length 1584;
Best Local Similarity 50.0%; Pred. No. 2.07e-01;
Matches 16; Conservative 3; Mismatches 12; Indels 1; Gaps 1;
Db 12 WILAPILLILLIGRRARAAAGADAGPPEPC 43
QY 10 WAAVTLILLILL-PPALLSSGAAQPLPDC 40

RESULT 8
ID O42394 PRELIMINARY; PRT; 205 AA.
AC O42394;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE NISHED (FRAGMENT).
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC GALLIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-SKELETAL MUSCLE;
RA DHAR M., MASCARENO E., SIDDIQUI M.A.Q.;
SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
EMBL; AF003093; G2232184;
NON_TER 1
SQ SEQUENCE 205 AA; 20829 MW; 57AA0342 CRC32;

Query Match 10.3%; Score 96; DB 12; Length 205;
Best Local Similarity 38.0%; Pred. No. 8.25e-01;
Matches 19; Conservative 15; Mismatches 13; Indels 3; Gaps 3;
Db 157 KNLPGRAGL-GR-EQFLCPAARGAGIPQIGHRASGTPRPSPALRASSCS 204
QY 68 KRSGPPGLQRLRLQASGNAAGILTMGRAGAPRCL-GRCS 116

RESULT 9
ID O06319 PRELIMINARY; PRT; 226 AA.
AC O06319;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JUL-1997 (TREMBLREL. 04, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 23.1 KD PROTEIN.
GN MTCY13E12.05.
OS MYCOBACTERIUM TUBERCULOSIS.
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; MYCOBACTERIACEAE.
RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA DEVLIN K., CHURCHER C.M.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RA BARRELL B.G., RAJANDREAM M.A., PARKHILL J.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE; 96181548.
RA PHILIPP W.J., POULET S., EIGMEIER K., PASCOPELLA L.,
RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
RA COLE S.T.;
RL PROC. NATL. ACAD. SCI. U.S.A. 93:3132-3137(1996).
DR EMBL; Z95390; E316057;
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 226 AA; 23113 MW; C7D17505 CRC32;

Query Match 10.2%; Score 95; DB 9; Length 226;
Best Local Similarity 27.3%; Pred. No. 1.08e+00;
Matches 21; Conservative 24; Mismatches 30; Indels 2; Gaps 2;

Db 96 FLAAGDANDASDHIOQASACRATRLVLGGYSGAAVIDIVTA-ADLPGL-GTQPLPP 153
QY 26 LLSGAAQPLPDCCKOKTSCRLYELLHGAGNAAGILTKRRSGPPGLQRLRLQ 85
Db 154 AADHTAAIALFGNPSG 170
QY 86 ASGNAAGILTMGRAG 102

RESULT 10
ID O15255 PRELIMINARY; PRT; 1337 AA.
AC Q15255;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE PROTEIN-TYROSINE PHOSPHATASE ETA PRECURSOR (EC 3.1.3.48) (R-PTD-ETA).
OS HOMO SAPIENS (HUMAN)
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95086212.
RA HONDA H., INAZAWA J., NISHIDA J., YAZAKI Y., HIRAI H.;
RL BLOOD 84:4186-4194(1994).
DR EMBL; D37781; G633073;
KW SIGNAL; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT; HYDROLASE.
FT SIGNAL 1 35 POTENTIAL.
FT CHAIN 36 1337 PROTEIN-TYROSINE PHOSPHATASE ETA.
FT DOMAIN 36 975 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 976 996 POTENTIAL.
FT DOMAIN 997 1337 CYTOPLASMIC (POTENTIAL).
FT DOMAIN ? ? ? FIBRONECTIN TYPE-III.
FT DOMAIN ? ? ? FIBRONECTIN TYPE-III.
FT DOMAIN ? ? ? FIBRONECTIN TYPE-III.
FT DOMAIN ? ? ? FIBRONECTIN TYPE-III.
FT DOMAIN ? ? ? FIBRONECTIN TYPE-III.
FT DOMAIN ? ? ? FIBRONECTIN TYPE-III.
FT DOMAIN ? ? ? FIBRONECTIN TYPE-III.
FT DOMAIN ? ? ? FIBRONECTIN TYPE-III.
FT ACT_SITE 1239 1239 BY SIMILARITY.
FT CARBOHYD 72 72 POTENTIAL.
FT CARBOHYD 82 82 POTENTIAL.
FT CARBOHYD 93 93 POTENTIAL.
FT CARBOHYD 104 104 POTENTIAL.
FT CARBOHYD 142 142 POTENTIAL.
FT CARBOHYD 172 172 POTENTIAL.

FT CHAIN 30 399 SEX HORMONE-BINDING GLOBULIN.
SQ SEQUENCE 399 AA; 43842 MW; 327FA1A7 CRC32;

Query Match 9.9%; Score 92; DB 10; Length 399;
Best Local Similarity 33.3%; Pred. No. 2.42e+00;
Matches 19; Conservative 14; Mismatches 20; Indels 4; Gaps 4;

Db 6 SVASLLLLLLPPHPTHGQVLRHVVPQTQNSQDSPA-R-Y-LSNGPGQEPVAVMTI 59
QY 11 AAVTLLLLLLPALSSGAAA-QPLPDCRCROKTCRLYELLHGAGNHAAGILTL 66

RESULT 15

ID Q99523 PRELIMINARY; PRT; 833 AA.

AC Q99523;

DT 01-MAY-1997 (TREMBLREL. 03, CREATED)

DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)

DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)

DE SORTILIN PRECURSOR.

GN SORTI.

OC HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

RN EUTHERIA; PRIMATES.

RP [1]

RP SEQUENCE FROM N.A.

RA PETERSEN C.M., NIELSEN M.S., JACOBSEN L., TOMMERUP N., HOLM H.,

RA ROIGARD H., GLIEMANN J., MADSEN P., MOESTRUP S.K.;

RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL; X98248; E246784; -

KW SIGNAL.

FT SIGNAL 1 33 POTENTIAL.

FT CHAIN 34 833 SORTILIN.

SQ SEQUENCE 833 AA; 92408 MW; D2E351B9 CRC32;

Query Match

Best Local Similarity 9.9%; Score 92; DB 2; Length 833;

Matches 14; Conservative 4; Mismatches 10; Indels 1; Gaps 1;

Db 14 WPHGLGLLLQLLPSTLSQDRIDAPP 42

QY 10 WA-AVTLLLLLLPALSSGAAQPLP 37

Search completed: Thu Jul 30 09:04:56 1998

Job time : 50 secs.

MPERCH_PP

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPERCH_PP protein - protein database search, using Smith-Waterman algorithm

on: Thu Jul 30 09:03:31 1998; MasPar time 5.27 Seconds
Regular output not generated. 623.327 Million cell updates/sec

Title: >US-08-938-548A-2
Description: (1-131) from US08938548A.pep
Perfect Score: 931
Sequence: 1 MNLPTKVSNAATVILLLLL.....GRCSPAPAAASVAPGGQSGI 131

Scoring table: PAM 150
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 41.687; Variance 80.083; scale 0.521

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	107	11.5	131	1	SECR_PIG SECRETIN PRECURSOR (FR	4.73e-03
2	102	11.0	1744	1	TENSIN	2.43e-02
3	101	10.8	599	1	PGH1_HUMAN PROTAGLANDIN G/H SYNT	3.36e-02
4	99	10.6	1013	1	PTPX_MACNE PROTEIN-TYROSINE PHOSP	6.37e-02
5	99	10.6	1015	1	PTPX_HUMAN PROTEIN-TYROSINE PHOSP	6.37e-02
6	98	10.5	1165	1	CYA6_MOUSE ADENYLATE CYCLASE, TYP	8.74e-02
7	97	10.4	260	1	URK1_MOUSE URIDINE KINASE (EC 2.7	1.20e-01
8	97	10.4	602	1	PGH1_MOUSE PROTAGLANDIN G/H SYNT	1.20e-01
9	96	10.3	205	1	YK07_YEAST HYPOTHETICAL 22.7 KD P	1.64e-01
10	94	10.1	101	1	GRO_CRIGR GROWTH REGULATED PROTE	3.04e-01
11	94	10.1	232	1	Y152_HUMAN HYPOTHETICAL PROTEIN K	3.04e-01
12	94	10.1	1337	1	PTPX_HUMAN PROTEIN-TYROSINE PHOSP	3.04e-01
13	94	10.1	2499	1	MPRI_BOVIN CATION-INDEPENDENT MAN	3.04e-01
14	92	9.9	235	1	FL3L_HUMAN SL CYTOKINE PRECURSOR	5.60e-01
15	92	9.9	251	1	C10B_HUMAN COMPLEMENT C1Q SUBCOMP	5.60e-01
16	92	9.9	676	1	ICP0_HSVBJ TRANS-ACTING TRANSCRIP	5.60e-01
17	92	9.9	676	1	ICP0_HSVBK TRANS-ACTING TRANSCRIP	5.60e-01
18	91	9.8	492	1	COG1_MOUSE STROMELYSIN-3 PRECURSO	1.37e-01
19	89	9.6	76	1	CD24_MOUSE SIGNAL TRANSDUCER CD24	1.37e-01
20	89	9.6	90	1	E PROTEIN	1.37e-01
21	89	9.6	90	1	VGE_BPHX3 LYSIS PROTEIN	1.37e-01
22	89	9.6	232	1	FL3L_MOUSE SL CYTOKINE PRECURSOR	1.37e-01
23	89	9.6	322	1	YCEC_HAEIN HYPOTHETICAL PROTEIN H	1.37e-01

24	89	9.6	430	1	SECR_PIG SECRETIN PRECURSOR (FR	4.73e-03
25	89	9.6	438	1	LCAT_MOUSE PHOSPHATIDYLCHOLINE-ST	1.37e-00
26	88	9.5	238	1	EPH3_HUMAN EPHRIN-A3 PRECURSOR (E	1.85e-00
27	88	9.5	317	1	LIP1_PSYIM LIPASE 1 PRECURSOR (EC	1.85e-00
28	88	9.5	1165	1	CYA6_MOUSE ADENYLATE CYCLASE, TYP	1.85e-00
29	88	9.5	1166	1	CYA6_MOUSE ADENYLATE CYCLASE, TYP	1.85e-00
30	87	9.3	229	1	PRL_FELCA PROLACTIN PRECURSOR (P	2.47e-00
31	87	9.3	245	1	ICP3_HSVIN INFECTED CELL PROTEIN	2.47e-00
32	87	9.3	319	1	YCEC_ECOLI HYPOTHETICAL 36.0 KD P	2.47e-00
33	87	9.3	419	1	HFLK_SHEEP HFLK PROTEIN	2.47e-00
34	87	9.3	600	1	PGH1_SHEEP PROTAGLANDIN G/H SYNT	2.47e-00
35	87	9.3	696	1	LSHR_PIG LUTROPIN-CHORIOGONADOT	2.47e-00
36	87	9.3	2491	1	MPRI_HUMAN CATION-INDEPENDENT MAN	2.47e-00
37	86	9.2	254	1	41BL_HUMAN 4-1BB LIGAND (4-1BBL)	3.30e-00
38	86	9.2	286	1	SSRA_RABIT TRANSLOCIN-ASSOCIATED	3.30e-00
39	86	9.2	370	1	WNT1_MOUSE WNT-1 PROTO-ONCOGENE P	3.30e-00
40	86	9.2	1027	1	CAFF_RIFPA FIBRIL-FORMING COLLAG	3.30e-00
41	86	9.2	1061	1	ANPA_HUMAN ATRIAL NATRIURETIC PEP	3.30e-00
42	86	9.2	1103	1	CYGD_HUMAN RETINAL GUANYLYL CYCLA	3.30e-00
43	86	9.2	2194	1	SC16_YEAST MULTIDOMAIN VESICLE CO	3.30e-00
44	85	9.1	171	1	YCX7_YEAST VERY HYPOTHETICAL 19.8	4.40e-00
45	85	9.1	325	1	NARG_MOUSE GPI-LINKED NAD(P)(+)--	4.40e-00

ALIGNMENTS

RESULT	ID	SECR_PIG	STANDARD;	PRT;	131 AA.
AC	P01279;				
DT	21-JUL-1986	(REL. 01, CREATED)			
DT	01-APR-1990	(REL. 14, LAST SEQUENCE UPDATE)			
DT	01-NOV-1995	(REL. 32, LAST ANNOTATION UPDATE)			
DE	SECRETIN PRECURSOR (FRAGMENT)				
GN	SCT.				
OS	SUS SCROFA (PIG), BOS TAURUS (BOVINE), AND CAVIA PORCELLUS (GUINEA				
OS	PIG)				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; ARTIODACTYLA.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES-PIG;				
RX	MEDLINE: 90192795.				
RA	KOPIN A.S., WHEELER M.B., LEITER A.B.;				
RL	PROC. NATL. ACAD. SCI. U.S.A. 87:2299-2303(1990).				
RN	[2]				
RP	SEQUENCE OF 30-56.				
RC	SPECIES-PIG;				
RX	MEDLINE: 70282334.				
RA	MUTT V., JORPES J.E., MAGNUSSON S.;				
RL	EUR. J. BIOCHEM. 15:513-519(1970).				
RN	[3]				
RP	SEQUENCE OF 30-59 AND 92-131.				
RC	SPECIES-PIG;				
RX	MEDLINE: 90370867.				
RA	GAFVELIN G., JOERNVALL H., MUTT V.;				
RL	PROC. NATL. ACAD. SCI. U.S.A. 87:6781-6785(1990).				
RN	[4]				
RP	SYNTHESIS OF 30-131.				
RC	SPECIES-PIG;				
RA	BODANSZKY M., ONDETTI M.A., LEVINE S.D., NARAYANAN V.L., SALTZA M.V.,				
RL	SHEEHAN J.T., WILLIAMS N.J., SABO E.F.;				
RN	[5]				
RP	SEQUENCE OF 30-56.				
RC	SPECIES-BOVINE;				
RX	MEDLINE: 81237102.				
RA	CARLOQUIST M., JOERNVALL H., MUTT V.;				
RL	FEBS LETT. 127:71-74(1981).				
RN	[6]				
RP	SEQUENCE OF 30-56.				
RC	SPECIES-C.PORCELLUS;				
RX	MEDLINE: 90254163.				
RA	BUSCAILL L., CAUVIN A., GOSSEN D., DE NEEF P., RATHE J.,				


```

01-NOV-1997 (REL. 35, CREATED)
01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
PROTEIN-TYROSINE PHOSPHATASE X PRECURSOR (EC 3.1.3.48) (R-PTP-X)
(M1851).
TPRN2.
GN
MACACA NEMESTRINA (PIG-TAILED MACAQUE).
OC
EUKARYOTA; METAZOA; CHORDATA; TETRAPODA; MAMMALIA;
OC
EUTHERIA; PRIMATES.
RN
[1]
RP
SEQUENCE FROM N.A.
RC
TISSUE=PANCREAS;
RA
LAGASSE J.M., JELINEK L., SEXSON S., LOFTON-DAY K., BREININGER J. J.
R
SHEPPARD P., KINDSVOGEL W., HAGOPIAN W.A.;
RL
MOL. MED. 3:163-173(1997).
CC
-1- FUNCTION: IMPLICATED IN DEVELOPMENT OF NERVOUS SYSTEM AND
CC
PANCREATIC ENDOCRINE CELLS.
CC
-1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O =
CC
PROTEIN TYROSINE + ORTHOPHOSPHATE.
CC
-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (PROBABLE).
CC
-1- TISSUE SPECIFICITY: PANCREAS.
CC
-1- DISEASE: AUTOANTIGEN IN INSULIN-DEPENDENT DIABETES MELLITUS.
CC
-1- SIMILARITY: CONTAINS 1 PROTEIN-TYROSINE PHOSPHATASE DOMAIN.
DR
EMBL; U91574; G1916942; -.
DR
PROSITE; PS00383; TYR-PHOSPHATASE_1; 1.
DR
PROSITE; PS50056; TYR-PHOSPHATASE_2; 1.
DR
PROSITE; PS50055; TYR-PHOSPHATASE_PTP; 1.
KW
HYDROLASE; RECEPTOR; GLYCOPROTEIN; SIGNAL; TRANSMEMBRANE.
FT
SIGNAL 1 20
FT
CHAIN 21 1013
FT
DOMAIN 22 613
FT
TRANSMEM 614 634
FT
DOMAIN 635 1013
FT
ACT_SITE 943 943
FT
CARBOHYD 562 562
FT
SEQUENCE 1013 AA; 111190 MW; 98C7D263 CRC32;
SQ
Query Match 10.6%; Score 99; DB 1; Length 1013;
Best Local Similarity 57.1%; Pred. No. 6.37e-02; Indels 0; Gaps
Matches 12; Conservative 7; Mismatches 2;
Db 6 LLLLLLLPPRVLPAAPSSVP 26
QY LLLLLLLPPALLSSGAAAP 35
RESULT 5
ID PTPX_HUMAN STANDARD; PRT; 1015 AA.
AC Q92932; Q92662;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE DE
DE PROTEIN-TYROSINE PHOSPHATASE X PRECURSOR (EC 3.1.3.48) (R-PTP-X)
DE (ISLET CELL AUTOANTIGEN RELATED PROTEIN) (ICAR) (PHOGRIN).
GN TPRN2.
GN HOMO SAPIENS (HUMAN).
OC
EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC
EUTHERIA; PRIMATES.
RN
[1]
RP
SEQUENCE FROM N.A.
RC
TISSUE=PANCREAS;
RC
TISSUE=PANCREAS;
RX
MEDLINE: 97032784.
RA
KAWASAKI E., HUTTON J.C., EISENBARTH G.S.;
RL
BIOCHEM. BIOPHYS. RES. COMMUN. 227:440-447(1996).
RN
[2]
RN
SEQUENCE FROM N.A.
RC
TISSUE=BRIN;
RC
SMITH P.D., WANG J., BARKER K.T., CROMPTON M.;
RL
SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN
[3]
RN
SEQUENCE FROM N.A.
RX
MEDLINE: 96394649.
RX
CUI L., YU W.-P., DE AIZPURUA H.J., SCHMIDT R.S., PALLEN C.J.;

```

RESULT 4
ID PTPX_MACNE
AC 002695: STANDARD; PRT; 1013 AA.

ID PGH_MOUSE STANDARD; PRT; 602 AA.
 AC P22437;
 DT 01-AUG-1991 (REL. 19, CREATED)
 DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
 DE PROTAGLANDIN G/H SYNTHASE 1 (CYCLOOXYGENASE)
 DE -1) (COX-1) (PROTAGLANDIN-ENDORPEOXIDE SYNTHASE 1) (PROTAGLANDIN H2
 DE SYNTHASE 1) (PGH SYNTHASE 1) (PGHS-1) (PHS 1).
 GN PTGS1 OR COX1 OR COX-1.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 90203007.
 RA DEWITT D.L., EL-HARITH E.A., KRAEMER S.A., ANDREWS M.J., YAO E.F.,
 RA ARMSTRONG R.L., SMITH W.L.;
 RJ J. BIOL. CHEM. 265:5192-5198(1990).
 CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN REGULATING OR PROMOTING
 CELL PROLIFERATION IN SOME NORMAL AND NEOPLASTICALLY TRANSFORMED
 CELLS.
 CC -!- CATALYTIC ACTIVITY: ARACHIDONATE + AH(2) + 2 O(2) -> PROTAGLANDIN
 H2 + A + H(2)O.
 CC -!- PATHWAY: FIRST STEP IN THE FORMATION OF PROTAGLANDINS AND
 THROMBOXANES.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED. MICROSOAL MEMBRANE.
 CC -!- THIS ENZYME ACTS BOTH AS A DIOXYGENASE AND AS A PEROXIDASE.
 CC -!- THIS ENZYME IS THE TARGET OF NONSTEROIDAL ANTI-INFLAMMATORY DRUGS
 SUCH AS ASPIRIN.
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
 DR EMBL: M34141; G200303;
 DR PIR: A35564; A35564.
 DR HSSP: P01132; 1EGF.
 DR MGD; MG1:97797; PTGS1.
 DR PROSITE: PS00022; EGF 1; FALSE NEG.
 DR PROSITE: PS01186; EGF 2; FALSE NEG.
 KW OXIDOREDUCTASE; DIOXYGENASE; PEROXIDASE; GLYCOPROTEIN; ACETYLATION;
 KW PROTAGLANDIN BIOSYNTHESIS; HEME; IRON; SIGNAL; MEMBRANE;
 KW EGF-LIKE DOMAIN.
 FT SIGNAL 1 26
 FT CHAIN 27 602 PROTAGLANDIN G/H SYNTHASE 1.
 FT DOMAIN 34 72 EGF-LIKE.
 FT ACT_SITE 209 209 DISTAL HISTIDINE (BY SIMILARITY).
 FT ACT_SITE 387 387 CYCLOOXYGENASE (BY SIMILARITY).
 FT BINDING 330 330 PROXIMAL HEME LIGAND (BY SIMILARITY).
 FT MOD_RES 532 532 ASPIRIN-ACETYLATED SERINE.
 FT DISULFID 38 49 BY SIMILARITY.
 FT DISULFID 43 59 BY SIMILARITY.
 FT DISULFID 61 71 BY SIMILARITY.
 FT DISULFID 39 161 BY SIMILARITY.
 FT DISULFID 571 577 BY SIMILARITY.
 FT CARBOHYD 70 70 POTENTIAL.
 FT CARBOHYD 106 106 POTENTIAL.
 FT CARBOHYD 146 146 POTENTIAL.
 SQ SEQUENCE 602 AA; 69042 MW; 96489281 CRC32;
 Query Match 10.4%; Score 97; DB 1; Length 602;
 Best Local Similarity 38.7%; Pred. No. 1-20e-01;
 Matches 12; Conservative 7; Mismatches 12; Indels 0; Gaps 0;
 Db 9 WFLPLLLLPPTSPVLLADPGVPSPVNPCC 39
 QY 10 WAAVTLTLLLLPPALLSSGAAQPLPDC 40
 RESULT 9
 ID YK07_YEAST STANDARD; PRT; 205 AA.
 AC P36061;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 22.7 KD PROTEIN IN SDH1-C1M5/YTA3 INTERGENIC REGION.

GN YKL147C OR YKL601.
 OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN-S288C;
 RA VANDENBOL M., BOLLE P.-A., DION C., PORTELETTE D., HILGER F.;
 RA YEAST 10:S35-S40(1994).
 DR EMBL: Z26877; G407501;
 DR EMBL: Z28146; G486251;
 DR PIR: S37804; S37804.
 DR PIR: S44581; S44581.
 KW HYPOTHETICAL PROTEIN.
 SQ SEQUENCE 205 AA; 22673 MW; 0AE8AD4E CRC32;
 Query Match 10.3%; Score 96; DB 1; Length 205;
 Best Local Similarity 53.8%; Pred. No. 1.64e-01;
 Matches 14; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
 Db 74 LLLLLLLLLPLPLPSVKGEPDAC 99
 QY 15 LLLLLLLLLPALLSSGAAQPLPDC 40
 RESULT 10
 ID GRO_CRIGR STANDARD; PRT; 101 AA.
 AC P09340;
 DT 01-MAR-1989 (REL. 10, CREATED)
 DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
 DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
 DE GROWTH REGULATED PROTEIN PRECURSOR.
 DE GRO.
 GN CRICETULUS GRISEUS (CHINESE HAMSTER).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 88041072.
 RA ANISOWICZ A., BARDWELL L., SAGER R.;
 RJ PROC. NATL. ACAD. SCI. U.S.A. 84:7188-7192(1987).
 CC -!- FUNCTION: HAS CHEMOTACTIC ACTIVITY FOR NEUTROPHILS.
 CC -!- SIMILARITY: BELONGS TO THE INTERCRINE ALPHA FAMILY (SMALL CYTOKINE
 C-X-C) (CHEMOKINE CXCL).
 CC EMBL: J03560; G304509;
 DR PIR: B28414; B28414.
 DR HSSP: P09341; 1MGs.
 DR PROSITE: PS00471; SMALL_CYTOKINES_CXC; 1.
 KW CYTOKINE; GROWTH FACTOR; INFLAMMATORY RESPONSE; SIGNAL.
 FT SIGNAL 1 28 POTENTIAL.
 FT CHAIN 29 101 GRO PROTEIN.
 FT DISULFID 37 63 BY SIMILARITY.
 FT DISULFID 39 79 BY SIMILARITY.
 SQ SEQUENCE 101 AA; 10893 MW; 3F83AD41 CRC32;
 Query Match 10.1%; Score 94; DB 1; Length 101;
 Best Local Similarity 41.7%; Pred. No. 3.04e-01;
 Matches 20; Conservative 9; Mismatches 15; Indels 4; Gaps 4;
 Db 3 PATR-SLLRAPLLLLLLLSLATSRLATGAPVANEILR-CQCLQTGTGVHL 48
 QY 4 PSTKVSAAVTLTLLLLPPALLSSGAA-AQPLPDC-CRQKTCSCRL 49
 RESULT 11
 ID Y152_HUMAN STANDARD; PRT; 292 AA.
 AC Q14165;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL PROTEIN KIA0152.
 GN KIA0152.
 OS HOMO SAPIENS (HUMAN).

OF THE COMPLEX. THIS RECEPTOR ALSO BINDS INSULIN GROWTH FACTOR II.
-!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. LYSOSOMAL.
-!- DOMAIN: CONTAINS 15 REPEATING UNITS OF APPROXIMATELY 147 AA. THE
MOST HIGHLY CONSERVED REGION WITHIN THE REPEAT CONSISTS OF A
STRETCH OF 13 AA THAT CONTAINS CYSTEINES AT BOTH ENDS.

EMBL: J03527; G162874; -;
PIR: A30788; A30788;
HSSP: P02784; LPDC.
PROSITE: PS00023; FIBRONECTIN_2; 1.
TRANSMEMBRANE; TRANSPORT; GLYCOPROTEIN; REPEAT; RECEPTOR; LYSOSOME;
SIGNAL.

1 44
FT SIGNAL CATION-INDEPENDENT MANNOSE-6-PHOSPHATE
FT CHAIN 45 2499
RECEPTOR.

FT DOMAIN 45 2313
FT TRANSMEM 2314 2336
FT DOMAIN 2337 2499
CYTOPLASMIC (POTENTIAL).

FT REPEAT 45 170
1.
FT REPEAT 171 327
2.
FT REPEAT 328 478
3.
FT REPEAT 479 629
4.
FT REPEAT 630 771
5.
FT REPEAT 772 933
6.
FT REPEAT 934 1089
7.
FT REPEAT 1090 1229
8.
FT REPEAT 1230 1373
9.
FT REPEAT 1374 1518
10.
FT REPEAT 1519 1658
11.
FT REPEAT 1659 1807
12.
FT REPEAT 1808 1959
13.
FT REPEAT 1960 2137
14.
FT REPEAT 2138 2290
15.
FT DOMAIN 120 120
FIBRONECTIN TYPE-II.

FT CARBOHYD 409 420
POTENTIAL.

FT CARBOHYD 444 444
POTENTIAL.

FT CARBOHYD 552 552
POTENTIAL.

FT CARBOHYD 590 590
POTENTIAL.

FT CARBOHYD 635 635
POTENTIAL.

FT CARBOHYD 755 755
POTENTIAL.

FT CARBOHYD 879 879
POTENTIAL.

FT CARBOHYD 959 959
POTENTIAL.

FT CARBOHYD 1030 1030
POTENTIAL.

FT CARBOHYD 1173 1173
POTENTIAL.

FT CARBOHYD 1255 1255
POTENTIAL.

FT CARBOHYD 1321 1321
POTENTIAL.

FT CARBOHYD 1665 1665
POTENTIAL.

FT CARBOHYD 1766 1766
POTENTIAL.

FT CARBOHYD 1825 1825
POTENTIAL.

FT CARBOHYD 2094 2094
POTENTIAL.

FT CARBOHYD 2145 2145
POTENTIAL.

FT CARBOHYD 2220 2220
POTENTIAL.

FT SEQUENCE 2499 AA; 274526 MW; 1AF70E56 CRC32;

Query Match 10.1%; Score 94; DB 1; Length 2499;

Best Local Similarity 60.0%; Pred. No. 3.04e-01;

Matches 12; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Db 30 LLLLLLLPPGWVGAAGTQ 49

QY 15 LLLLLLLPPALLSSGAAQ 34

RESULT 14

ID FL3L HUMAN STANDARD; PRT; 235 AA.

AC P49771;

DT 01-OCT-1996 (REL. 34, CREATED)

DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)

DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE SL CYTOKINE PRECURSOR (FLT3 LIGAND).

GN FL3LG.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; PRIMATES.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 94193428.

RA HANNUM C., CULPEPPER J., CAMPBELL D., MCCLANAHAN T., ZURAWSKI S.,

RA BAZAN J.F., KASTLEIN R., HUDAK S., WAGNER J., MATTSON J., LUH J.,

RA DUDA G., MARTINA N., PETERSON D., MENON S., SHANAFELT A.,

RA MUENCH M., KELNER G., NAMIKAWA R., RENNICK D., RONCAROLO M.G.,

RA ZLOTNIK A., ROSNET O., DUBREUIL P., BIRNBAUM D., LEE F.,

RL NATURE 368:643-648(1994).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE; 94235842.

RA LYMAN S.D., JAMES L., JOHNSON L., BRASEL K., DE VRIES P.,

RA ESCOBAR S.S., DOWNEY H., SPLETT R.R., BECKMANN M.P., MCKENNA H.J.,

RL BLOOD 83:2795-2801(1994).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE; 96032581.

RA LYMAN S.D., STOCKING K., DAVISON B., FLETCHER F., JOHNSON L.,

RA ESCOBAR S.,

RL ONCOGENE 11:1165-1172(1995).

CC -!- FUNCTION: STIMULATES THE PROLIFERATION OF EARLY HEMATOPOIETIC

CELLS. SYNERGIZES WELL WITH A NUMBER OF OTHER COLONY STIMULATING

FACTORS AND INTERLEUKINS.

CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. A SOLUBLE FORM

IS ALSO PRODUCED BY ALTERNATIVE SPLICING.

CC -!- ALTERNATIVE PRODUCTS: VARIOUS FORMS ARE PRODUCED BY ALTERNATIVE

SPLICING.

CC EMBL; U04806; G43845; -;

DR EMBL; U03858; G494979; -;

DR EMBL; U29874; G1072037; -;

DR EMBL; U29874; G1072038; -;

DR MIM; 600007; -;

KW CYTOKINE; GLYCOPROTEIN; TRANSMEMBRANE; ALTERNATIVE SPLICING; SIGNAL.

FT SIGNAL 1 26

FT CHAIN 27 235

FT DOMAIN 27 184

FT TRANSMEM 185 205

FT DOMAIN 206 235

FT CARBOHYD 126 126

FT CARBOHYD 149 149

FT VARSPLIC 161 178

FT VARSPLIC 179 235

FT CONFLICT 72 72

FT SEQUENCE 235 AA; 26416 MW; F3F8AB35 CRC32;

Query Match 9.9%; Score 92; DB 1; Length 235;

Best Local Similarity 55.6%; Pred. No. 5.60e-01;

Matches 10; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Db 7 AWSPTTLLLLLSGL 24

QY 9 SWAAVTLTLLLLLPAL 26

RESULT 15

ID CLOB HUMAN STANDARD; PRT; 251 AA.

AC P02746;

DT 21-JUL-1986 (REL. 01, CREATED)

DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)

DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)

DE COMPLEMENT C1Q SUBCOMPONENT, B CHAIN PRECURSOR.

GN CLOB.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; PRIMATES.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 86076906.

RA REID K.B.M.;

RL BIOCHEM. J. 231:729-735(1985).

WIREH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd
MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Jul 30 09:02:52 1998; MasPar time 7.57 Seconds
632.140 Million cell updates/sec
Linear output not generated.

Title: >US-08-938-548A-2
Description: (1-131) from US08938548A.pep
Perfect Score: 931
Sequence: 1 MNLPSKVSNAVTLILL.....GRRCSAPAAASVAPGGSGI 131
Scoring table: PAM 150
Gap 11
Searched: 120441 seqs, 36531193 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: pir56
1:pir1 2:pir2 3:pir3 4:pir4 5:nrl3d
Statistics: Mean 39.642; Variance 90.062; scale 0.440

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description	Pred. No.
1	109	11.7	602	2	prostaglandin G/H syn	1.59e-02
2	108	11.6	599	2	prostaglandin-endoper	2.13e-02
3	107	11.5	131	1	secretin precursor -	2.85e-02
4	107	11.5	602	2	cyclooxygenase 1 - ra	2.85e-02
5	105	11.3	347	2	hypothetical protein	5.09e-02
6	103	11.1	1792	2	tensin - chicken (fra	9.04e-02
7	102	11.0	1733	2	tensin - chicken	1.20e-01
8	102	11.0	1744	2	tensin, cardiac muscl	1.20e-01
9	101	10.8	562	2	prostaglandin-endoper	1.60e-01
10	101	10.8	599	2	prostaglandin-endoper	1.60e-01
11	99	10.6	245	2	phogrin - human	2.80e-01
12	99	10.6	1015	2	transmembrane tyrosin	2.80e-01
13	99	10.6	1015	2	adenylate cyclase (EC	3.70e-01
14	98	10.5	1165	2	adenylate cyclase (EC	3.70e-01
15	98	10.5	1166	2	type 5 adenylyl cycla	4.88e-01
16	97	10.4	602	2	prostaglandin-endoper	6.43e-01
17	96	10.3	205	2	hypothetical protein	6.43e-01
18	96	10.3	312	2	hypothetical protein	6.43e-01
19	95	10.2	491	2	stromelysin 3 (EC 3.4	8.46e-01
20	94	10.1	101	2	growth-regulated prot	1.11e+00
21	94	10.1	1337	2	protein-tyrosine phos	1.11e+00
22	94	10.1	1337	1	protein-tyrosine-phos	1.11e+00
23	94	10.1	2499	2	mannose 6-phosphate r	1.11e+00

24	92	9.9	235	2	S43292	FLT3/FLK2 ligand (clo	1.90e+00
25	92	9.9	253	1	C1HQB	complement subcompone	1.90e+00
26	92	9.9	676	1	EDBE23	immediate-early prote	1.90e+00
27	92	9.9	676	1	EDBE22	immediate-early prote	1.90e+00
28	91	9.8	492	2	A44399	stromelysin 3 (EC 3.4	3.24e+00
29	90	9.7	147	2	S24303	hypothetical protein	4.22e+00
30	89	9.6	76	2	A43537	heat-stable antigen M	4.22e+00
31	89	9.6	90	1	ZEBPF4	gene E protein - phag	4.22e+00
32	89	9.6	90	2	ZEBPF4	gene E protein - phag	4.22e+00
33	89	9.6	91	2	J50455	lysis protein - phage	4.22e+00
34	89	9.6	220	2	I58343	flt3 ligand isoform 5	4.22e+00
35	89	9.6	220	2	S43291	FLT3/FLK2 ligand (clo	4.22e+00
36	89	9.6	231	2	A49265	FLT3/FLK-2 ligand pre	4.22e+00
37	89	9.6	322	2	G64151	hypothetical protein	4.22e+00
38	89	9.6	430	2	A24702	serine proteinase sna	4.22e+00
39	89	9.6	438	1	XXMSN	phosphatidylcholine--	4.22e+00
40	88	9.5	169	2	S33141	hypothetical protein	5.48e+00
41	88	9.5	317	2	S28225	triacylglycerol lipas	5.48e+00
42	88	9.5	317	2	S57275	triacylglycerol lipas	5.48e+00
43	88	9.5	607	2	I37560	protein-tyrosine kina	5.48e+00
44	88	9.5	1165	2	A46180	adenylyl cyclase type	5.48e+00
45	88	9.5	1180	2	A47202	adenylyl cyclase (EC	5.48e+00

ALIGNMENTS

RESULT 1
ENTRY S69198 #type complete
TITLE prostaglandin G/H synthase 1 - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 10-Sep-1997
ACCESSIONS S69198; S69199
REFERENCE S69198
#authors Kitzler, J.W.
#submission submitted to the EMBL Data Library, December 1994
#accession S69198
#status preliminary
#molecule_type mRNA
#residues 1-602 #label KIT
#cross-references EMBL:U18060; NID:g603051; PID:g603052
REFERENCE S69199
#authors Kitzler, J.; Hill, E.; Hardman, R.; Reddy, N.; Philpot, R.; Eling, T.E.
#journal Arch. Biochem. Biophys. (1995) 316:856-863
#title Analysis and quantitation of splicing variants of the TPA-inducible PGHS-1 mRNA in rat tracheal epithelial cells.
#accession S69199
#molecule_type mRNA
#residues 61-602 #label KI2
#cross-references EMBL:U18060
#note only a part of the nucleic acid sequence is shown
KEYWORDS alternative splicing
SUMMARY #length 602 #molecular-weight 69032 #checksum 6994
Query Match 11.7%; Score 109; DB 2; Length 602;
Best Local Similarity 53.6%; Pred. No. 1.59e-02;
Matches 15; Conservative 6; Mismatches 5; Indels 2; Gaps 1;
Db 12 LLLLLLLLLPPVLLTDAGVSPVNPCC 39
QY 15 LLLLLLLLLPPA--LLSSGAAQPLPDC 40
RESULT 2
ENTRY A36746 #type complete
TITLE prostaglandin-endoperoxide synthase (EC 1.14.99.1) 1 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 30-Apr-1991 #sequence_revision 30-Apr-1991 #text_change 29-Aug-1997
ACCESSIONS A36746
REFERENCE A36746
#authors Yokoyama, C.; Tanabe, T.

```

9
RESULT      #type complete
ENTRY      prostaglandin-endoperoxide synthase (EC 1.14.99.1) 2
TITLE      precursor - human
ALTERNATE_NAMES  cyclooxygenase; prostaglandin G/H synthase
ORGANISM   #formal_name Homo sapiens #common_name man
DATE       31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change
           20-Mar-1998
REFERENCES  A38146
AUTHORS     Diaz, A.; Reginato, A.M.; Jimenez, S.A.
JOURNAL     J. Biol. Chem. (1992) 267:10816-10822
#authors
#journal
#title      Alternative splicing of human prostaglandin G/H synthase mRNA
            and evidence of differential regulation of the resulting
            transcripts by transforming growth factor beta 1,
            interleukin 1 beta, and tumor necrosis factor alpha.
#cross-references MIMD:92268138
#accession   A38146
#molecule_type mRNA
#residues    1-562 #label DIA
##cross-references GB:S36219; NID:G249623; PTD:G249624

```

[illegible]

MSRFLH

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:05:14 1998; MasPar time 1.81 Seconds
424.456 Million cell updates/sec
ular output not generated.

Title: >US-08-938-548A-2
Description: (1-131) from US08938548A.pep
Perfect Score: 931
Sequence: 1 MNLPTKVSNAATVILLLLL.....GRRCSAPAAASVAPGGSGI 131

Scoring table: PAM 150
Gap 11

Searched: 63816 seqs, 5850866 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5_COMB 2:PCP9_COMB 3:backfiles

Statistics: Mean 27.627; Variance 133.583; scale 0.207

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	94	10.1	1337	2	PCT-US95-0 Sequence 2, Applicatio	4.75e+00
2	92	9.9	235	1	US-08-243- Sequence 6, Applicatio	6.53e+00
3	92	9.9	235	2	PCT-US94-0 Sequence 6, Applicatio	6.53e+00
4	91	9.8	492	1	US-08-001- Sequence 4, Applicatio	7.65e+00
5	91	9.8	492	1	US-07-794- Sequence 4, Applicatio	7.65e+00
6	89	9.6	231	2	PCT-US94-0 Sequence 2, Applicatio	1.05e+01
7	89	9.6	231	1	US-08-220- Sequence 7, Applicatio	1.05e+01
8	89	9.6	231	1	US-08-243- Sequence 6, Applicatio	1.05e+01
9	89	9.6	231	2	PCT-US95-0 Sequence 2, Applicatio	1.05e+01
10	88	9.5	238	1	US-08-240- Sequence 2, Applicatio	1.23e+01
11	88	9.5	238	1	US-08-453- Sequence 2, Applicatio	1.23e+01
12	88	9.5	1165	1	US-08-240- Sequence 2, Applicatio	1.23e+01
13	86	9.2	254	1	US-08-236- Sequence 4, Applicatio	1.68e+01
14	85	9.1	422	1	US-08-469- Sequence 170, Applicat	1.96e+01
15	85	9.1	422	1	US-08-249- Sequence 170, Applicat	1.96e+01
16	85	9.1	422	1	US-08-428- Sequence 3, Applicatio	1.96e+01
17	85	9.1	422	2	PCT-US95-0 Sequence 170, Applicat	1.96e+01
18	85	9.1	422	2	PCT-US94-0 Sequence 166, Applicat	1.96e+01
19	85	9.1	422	2	PCT-US94-0 Sequence 185, Applicat	1.96e+01
20	85	9.1	422	1	US-08-036- Sequence 170, Applicat	1.96e+01
21	85	9.1	426	1	US-07-918- Sequence 2, Applicatio	1.96e+01
22	84	9.0	263	1	US-08-300- Sequence 2, Applicatio	2.29e+01
23	84	9.0	291	2	PCT-US95-0 Sequence 7, Applicatio	2.29e+01

24	82	8.8	291	3	5212074-5	Patent No. 5212074.	3.11e+01
25	82	8.8	292	3	5258287-24	Patent No. 5258287.	3.11e+01
26	82	8.8	1239	1	US-08-026-	Sequence 3, Applicatio	3.11e+01
27	81	8.7	171	1	US-08-313-	Sequence 4, Applicatio	3.62e+01
28	81	8.7	533	1	US-08-445-	Sequence 10, Applicati	3.62e+01
29	81	8.7	533	1	US-08-484-	Sequence 13, Applicati	3.62e+01
30	80	8.6	206	2	PCT-US95-0	Sequence 2, Applicatio	4.22e+01
31	80	8.6	206	1	US-08-197-	Sequence 2, Applicatio	4.22e+01
32	80	8.6	263	2	PCT-US91-0	Sequence 2, Applicatio	4.22e+01
33	80	8.6	361	1	US-08-415-	Sequence 7, Applicatio	4.22e+01
34	80	8.6	488	1	US-08-001-	Sequence 2, Applicatio	4.22e+01
35	80	8.6	488	1	US-07-794-	Sequence 2, Applicatio	4.22e+01
36	79	8.5	228	1	US-08-379-	Sequence 5, Applicatio	4.91e+01
37	79	8.5	242	1	US-08-289-	Sequence 6, Applicatio	4.91e+01
38	78	8.4	63	3	5304637-7	Patent No. 5304637.	5.71e+01
39	78	8.4	376	1	US-08-253-	Sequence 33, Applicati	5.71e+01
40	78	8.4	447	1	US-07-937-	Sequence 29, Applicati	5.71e+01
41	78	8.4	689	1	US-08-059-	Sequence 5, Applicatio	5.71e+01
42	78	8.4	689	2	PCT-US91-0	Sequence 5, Applicatio	5.71e+01
43	78	8.4	689	1	US-07-766-	Sequence 5, Applicatio	5.71e+01
44	77	8.3	831	1	US-08-483-	Sequence 5, Applicatio	6.64e+01
45	77	8.3	831	1	US-08-254-	Sequence 5, Applicatio	6.64e+01

ALIGNMENTS

RESULT 1
ID PCT-US95-05512-2 STANDARD; PRT; 1337 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 2, Application PC/TUS9505512
XX
CC Sequence 2, Application PC/TUS9505512
CC GENERAL INFORMATION:
CC APPLICANT: Tonks, Nicholas K. and stnan, Arne
CC TITLE OF INVENTION: Density Enhanced Protein Tyrosine
CC TITLE OF INVENTION: Phosphatase
CC NUMBER OF SEQUENCES: 6
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
CC ADDRESSEE: Borun
CC STREET: 233 South Wacker Drive, Suite 6300
CC CITY: Chicago
CC STATE: Illinois
CC COUNTRY: United States of America
CC ZIP: 60606
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/05512
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Borun, Michael F.
CC REGISTRATION NUMBER: 25,447
CC REFERENCE/DOCKET NUMBER: 27866/31954
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 312-474-6300
CC TELEFAX: 312-474-0448
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1337 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1337 AA; 146192 MW; 9857375 CN;

Query Match 10.1%; Score 94; DB 2; Length 1337;

Sequence 2, Application PC/TUS9405365

Sequence 2, Application US/08243545
 Patent No. 5554512
 GENERAL INFORMATION:
 APPLICANT: Lyman, Stewart D.
 APPLICANT: Beckmann, M. Patricia
 TITLE OF INVENTION: Ligands for flt3/flk-2 Receptors
 NUMBER OF SEQUENCES: 8
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Stephen L. Malaska, Immunex Corporation
 STREET: 51 University Street
 CITY: Seattle
 STATE: Washington
 COUNTRY: US
 ZIP: 98101
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: Apple Macintosh
 OPERATING SYSTEM: Macintosh 7.0.1
 SOFTWARE: Microsoft Word, Version #5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/243,545
 FILING DATE: 11-MAY-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/162,407
 FILING DATE: 03-DEC-1993
 APPLICATION NUMBER: 08/111,758
 FILING DATE: August 25, 1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/106,463
 FILING DATE: August 12, 1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/068,394
 FILING DATE: May 24, 1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Malaska, Stephen L.
 REGISTRATION NUMBER: 32,655
 REFERENCE/DOCKET NUMBER: 2813-C
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 587-0430
 TELEFAX: (206) 233-0644
 TELEX: 756822
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 231 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE 231 AA; 26027 MW; 272421 CN;

Query Match 9.6%; Score 89; DB 1; Length 231;
 Best Local Similarity 57.9%; Pred. No. 1.05e+01;
 Matches 11; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

Db 7 AWSPNSSLLLLLLLLSPCL 25
 :::: |||||:|
 QY 9 SWAA-VTLTLLLLLPPAL 26

RESULT 9
 ID PCT-US95-03866-6 STANDARD; PRT; 231 AA.
 AC xxxxxx
 XX
 DT
 DE Sequence 6, Application PC/TUS9503866
 CC Sequence 6, Application PC/TUS9503866
 CC GENERAL INFORMATION:
 APPLICANT: Beckmann, M. P.
 APPLICANT: CERRETTI, DOUGLAS P.
 TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: IMMUNEX CORPORATION

CC APPLICANT: CytoMed, Inc. (all states except US)
 CC APPLICANT: Nocka, Karl (US only)
 CC APPLICANT: Lobell, Robert B (US only)
 CC TITLE OF INVENTION: STABILIZED DIMER OF KIT LIGAND AND
 CC TITLE OF INVENTION: FLT-3/FLK-2 LIGAND
 CC NUMBER OF SEQUENCES: 36
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Fish & Neave
 CC STREET: 1251 Avenue of the Americas
 CC CITY: New York
 CC STATE: New York
 CC COUNTRY: United States of America
 CC ZIP: 10020
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: PCT/US95/03866
 CC FILING DATE:
 CC CLASSIFICATION:
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/220,379
 CC FILING DATE: 28-MAR-1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Haley Jr, James F
 CC REGISTRATION NUMBER: 27,794
 CC REFERENCE/DOCKET NUMBER: CytoMed/2
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 212-596-9000
 CC TELEFAX: 212-596-9090
 CC INFORMATION FOR SEQ ID NO: 6:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 231 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: not relevant
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC HYPOTHETICAL: NO
 CC FEATURE:
 CC NAME/KEY: Protein
 CC LOCATION: 1-205
 CC SEQUENCE 231 AA; 25999 MW; 274601 CN;

Query Match 9.6%; Score 89; DB 2; Length 231;
 Best Local Similarity 57.9%; Pred. No. 1.05e+01;
 Matches 11; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

Db 7 AWSPNSSLLLLLLLLSPCL 25
 :::: |||||:|
 QY 9 SWAA-VTLTLLLLLPPAL 26

RESULT 10
 ID US-08-240-124-2 STANDARD; PRT; 238 AA.
 AC xxxxxx
 XX
 DT
 DE Sequence 2, Application US/08240124
 CC Sequence 2, Application US/08240124
 CC Patent No. 5516658
 CC GENERAL INFORMATION:
 CC APPLICANT: BECKMANN, M. P.
 CC APPLICANT: CERRETTI, DOUGLAS P.
 CC TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE
 CC NUMBER OF SEQUENCES: 4
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: IMMUNEX CORPORATION

APPLICATION NUMBER: US/08/240,357
 FILING DATE: 10-MAY-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Gordon, Alan M.
 REGISTRATION NUMBER: 30,637
 REFERENCE/DOCKET NUMBER: 31,705-01
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 201-831-3244
 TELEFAX: 201-831-3305
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1165 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE 1165 AA; 130394 MW; 6766222 CN;
 Query Match 9.5%; Score 88; DB 1; Length 1165;
 Best Local Similarity 33.3%; Pred. No. 1.23e+01;
 Matches 15; Conservative 14; Mismatches 1; Indels 2; Gaps 2;
 DB 148 SSLTLLMAVLLVLTAVLLAFHAPARPOPAYVALLACATLEVAL 192
 QY 11 AAVTLL-LLLLPPALLS-SGAAQPLPDCCKQNTCSRLYELL 53
 RESULT 13
 ID US-08-236-918A-4 STANDARD; PRT; 254 AA.
 AC xxxxxx
 XX
 DT
 XX
 DE Sequence 4, Application US/08236918A
 XX Sequence 4, Application US/08236918A
 CC Patent No. 5674704
 CC GENERAL INFORMATION:
 CC APPLICANT: Alderson, Mark R.
 CC APPLICANT: Goodwin, Raymond G.
 CC APPLICANT: Smith, Craig A.
 CC TITLE OF INVENTION: Cytokine Designated 4-1BB Ligand
 CC NUMBER OF SEQUENCES: 18
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
 CC STREET: 51 University Street
 CC CITY: Seattle
 CC STATE: Washington
 CC COUNTRY: US
 CC ZIP: 98101
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: Apple Power Macintosh
 CC OPERATING SYSTEM: Apple 7.5.3
 CC SOFTWARE: Microsoft Word, Version #6.0.1
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/236,918A
 CC FILING DATE: 06-May-1994
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/060,843
 CC FILING DATE: 07-May-1993
 CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Anderson, Kathryn A.
 CC REGISTRATION NUMBER: 32,172
 CC REFERENCE/DOCKET NUMBER: 2801-B
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (206) 587-0430
 CC TELEFAX: (206) 233-0644
 CC INFORMATION FOR SEQ ID NO: 4:
 CC SEQUENCE CHARACTERISTICS:

LENGTH: 254 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE 254 AA; 26624 MW; 316872 CN;
 Query Match 9.2%; Score 86; DB 1; Length 254;
 Best Local Similarity 61.1%; Pred. No. 1.68e+01;
 Matches 11; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 DB 27 LPWALVAGLLLLLLLLAA 44
 QY 8 VSWAAVTLLLLLLLLPPA 25
 RESULT 14
 ID US-08-469-569-170 STANDARD; PRT; 422 AA.
 XX
 AC xxxxxx
 XX
 DT
 XX
 DE Sequence 170, Application US/08469569
 XX Sequence 170, Application US/08469569
 CC Patent No. 5606032
 CC GENERAL INFORMATION:
 CC APPLICANT: Goodearl, Andrew; Stroobant, Paul;
 CC APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchioni, Mark;
 CC APPLICANT: Chen, Maio Su; Hiles, Ian
 CC TITLE OF INVENTION: Glial Mitogenic Factors, Their
 CC TITLE OF INVENTION: Preparation and Use
 CC NUMBER OF SEQUENCES: 184
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Felfe & Lynch
 CC STREET: 805 Third Avenue
 CC CITY: New York City
 CC STATE: New York
 CC COUNTRY: USA
 CC ZIP: 10022
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
 CC COMPUTER: IBM
 CC OPERATING SYSTEM: PC-DOS
 CC SOFTWARE: Wordperfect
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/469,569
 CC FILING DATE: 06-JUN-1995
 CC CLASSIFICATION: 530
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/036,555
 CC FILING DATE: 24-MAR-1993
 CC APPLICATION NUMBER: 07/965,173
 CC FILING DATE: 23-OCT-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/940,389
 CC FILING DATE: 03-SEP-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/907,138
 CC FILING DATE: 30-JUN-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/863,703
 CC FILING DATE: 03-APRIL-1992
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: U.K. 91 07566.3
 CC FILING DATE: 10-APRIL-1991
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Tsai, Christine H.
 CC REGISTRATION NUMBER: 34,266
 CC REFERENCE/DOCKET NUMBER: LUD 5250.4
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (212) 688-9200
 CC TELEFAX: (212) 838-3884

WORLD

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

on: Thu Jul 30 09:02:10 1998; MasPar time 5.32 Seconds
377.841 Million cell updates/sec
Modular output not generated.

Title: >US-08-938-548A-2
Description: (1-131) from US08938548A.pep
Perfect Score: 931
Sequence: 1 MNLPTKVSNAVTLILLL.....GRCSPAASVAPGGQSGI 131

Scoring table: PAM 150
Gap 11

Searched: 124785 seqs, 15338987 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq31-2
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27

Statistics: Mean 29.691; Variance 135.127; scale 0.220

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	108	11.6	599	4	R21690 Prostaglandin endoper	1.04e+00
2	99	10.6	989	24	W25170 Human insulinoma-asso	4.77e+00
3	99	10.6	986	24	W25171 Human insulinoma-asso	4.77e+00
4	99	10.6	1012	26	W35296 Macaque islet cell an	4.77e+00
5	99	10.6	1015	27	W35345 Human protein tyrosin	4.77e+00
6	99	10.6	1015	23	W18092 Type I diabetes-assoc	4.77e+00
7	94	10.1	1337	14	R85203 huDEP-1	1.09e+01
8	92	9.3	235	12	R67941 Human flt-3 ligand.	1.52e+01
9	92	9.9	235	13	R66175 Human S86/S109 Flt3	1.52e+01
10	91	9.8	492	5	R24863 Sequence of pre-pro s	1.78e+01
11	89	9.6	35	1	P94256 Truncated E protein f	2.47e+01
12	89	9.6	231	12	R67540 Mouse flt-3 ligand.	2.47e+01
13	89	9.6	232	13	R66177 Mouse Mofl10/T118 Flt	2.47e+01
14	88	9.5	234	16	R82605 Eph transmembrane tyr	2.90e+01
15	88	9.5	238	13	R71481 Human hex-L protein.	2.90e+01
16	88	9.5	480	23	W22849 Osteoblast like cell	2.90e+01
17	88	9.5	1165	7	R37309 Cardiac adenylyl cycl	2.90e+01
18	87	9.3	186	6	R32428 Wheat germ agglutinin	3.40e+01
19	87	9.3	551	13	R77858 S. clavuligerus ORF1	3.40e+01

20	87	9.3	600	2	P91008 Prostaglandin endoper	3.40e+01
21	86	9.2	126	8	R50054 ICP34.5 fragment.	3.99e+01
22	86	9.2	254	25	W26657 Human 4-1BB ligand.	3.99e+01
23	86	9.2	254	12	R64190 Human 4-1BB-L polypep	3.99e+01
24	86	9.2	1481	19	W03740 rchd528 gene product.	3.99e+01
25	85	9.1	248	12	R67242 Human glial cell gro	4.68e+01
26	85	9.1	248	21	W09358 Human glial growth fa	4.68e+01
27	85	9.1	248	16	R87445 Human glial growth fa	4.68e+01
28	85	9.1	248	9	R55653 GGF segment E.	4.68e+01
29	85	9.1	248	17	R96074 Human glial growth fa	4.68e+01
30	85	9.1	248	9	R46912 GGF segment E.	4.68e+01
31	85	9.1	363	23	W12414 Porcine complement in	4.68e+01
32	85	9.1	422	12	R67258 Human glial cell grow	4.68e+01
33	85	9.1	422	21	W09371 Human neuregulin GGF2	4.68e+01
34	85	9.1	422	9	R46923 GGF-II encoded by clo	4.68e+01
35	85	9.1	422	16	R87467 Glial growth factor s	4.68e+01
36	85	9.1	422	16	R86628 Mature hGGF2.	4.68e+01
37	85	9.1	422	21	W09372 Human GGF2.	4.68e+01
38	85	9.1	422	16	R87466 Glial growth factor s	4.68e+01
39	85	9.1	422	17	R96081 Glial growth factor G	4.68e+01
40	85	9.1	422	9	R55654 GGF-II encoded by clo	4.68e+01
41	85	9.1	426	6	R31036 C. acidovorans stereo	4.68e+01
42	84	9.0	263	16	R90840 Murine interleukin-15	5.49e+01
43	84	9.0	291	15	R89952 Insulin-like growth f	5.49e+01
44	84	9.0	438	27	W37355 Wzy (Rfc) protein inv	5.49e+01
45	84	9.0	1865	19	W03515 Human DOCK180 protein	5.49e+01

ALIGNMENTS

RESULT 1
ID R21690 standard; Protein; 599 AA.
AC R21690; 23-JUL-1992 (first entry)
DE Prostaglandin endoperoxide synthase.
KW Human; hPES; screening; anti-inflammatory; antibody.
OS Homo sapiens.
PN J04045786-A.
PF 14-FEB-1992.
PD 13-JUN-1990; 152784.
PR 13-JUN-1990; JP-152784.
PA (MEIP) MEIJI MILK PRODS KK.
DR WPI: 92-101937/13.
DR N-PSDB; Q23001.
PT Human prostaglandin endo-peroxide synthase - for screening
PT antiinflammatory agents
PS Claim 1; Page 1; 10pp; Japanese.
CC The protein sequence of hPES was deduced from the cDNA sequence
CC obtd. by screening a human genomic library in EMEL3. hPES can be
CC used to screen anti-inflammatory agents. An anti-body against a
CC peptide specific to hPES can be made, and used for the determin-
CC ation of the protein.
SQ Sequence 599 AA;

Query Match 11.68; Score 108; DB 4; Length 599;
Best Local Similarity 55.6%; Pred. No. 1.04e+00;
Matches 15; Conservative 4; Mismatches 7; Indels 1; Gaps 1;
Db 10 llllllllpilpvladpgaptvnpcc 36
QY 15 LLLLLLLLLP-PALLSSGAAQPLPDC 40

RESULT 2
ID W25170 standard; Protein; 969 AA.
AC W25170;
DT 15-DEC-1997 (first entry)
DE Human insulinoma-associated antigen 2-beta (short version).
KW Insulinoma-associated antigen 2 beta; IA-2 beta; autoantigen;
KW insulin dependent diabetes; IDD; antibody; diagnosis; susceptibility;
KW type 1; type 2.
OS Homo sapiens.
PN W09707211-A1.

```

OS Homo sapiens.
FH Key Location/Qualifiers
FT Domain 1..614
FT /note="extracellular domain"
FT Domain 615..706
FT /note="transmembrane domain"
FT Domain 707..1082
FT /note="intracellular domain"
FT Misc_difference 294
FT /note="residue 294 is Glu in protein encoded by
FT a pancreas cDNA clone"
FT Misc_difference 914
FT /note="predicted phosphotyrosine contact site"
FT Misc_difference 947
FT /note="predicted phosphotyrosine contact site"
FT Binding_site 776..779
FT /note="predicted rim of binding pocket"
FT Binding_site 951
FT /note="predicted base of binding pocket"
FT Region 911..919
FT /note="predicted surface loop"
WO9736918-A1.
PD 09-OCT-1997.
PF 25-MAR-1997; U05284.
PR 25-MAR-1997; US-816962.
PR 29-MAR-1996; US-014267.
PA (SUGB-) SUGEN INC.
PI Plowman GD;
PI WPI; 97-503042/46.
DR N-PSDB; T95368.
DR Protein tyrosine phosphatase MOT12 - used to develop products for
PT the diagnosis and treatment of, e.g. cell proliferative disorders,
PT diabetes and immune disorders
PT Claim 9; Fig 1; 64pp; English.
CC This protein sequence comprises a novel human receptor-type
CC protein tyrosine phosphatase, designated MOT12, suggested to play
CC a role in the growth, differentiation and survival of neurons in
CC the adult. The amino acid sequence was deduced from a coding
CC sequence (see T95368) of overlapping cDNA clones isolated from
CC human caudate nucleus and pancreas cDNA libraries. MOT12
CC expression is highly restricted, being found only in adult brain
CC and pancreas in neurosecretory cell types. Polypeptides, including
CC those that comprise residues 1-614, 615-706 or 707-1082 of MOT12 or
CC comprise MOT12 lacking one or more of the extracellular,
CC transmembrane or intracellular domains, can be expressed in host
CC cells. The MOT12 polypeptides, nucleic acids encoding them, cells,
CC tissues, transgenic animals and antibodies can be used in methods
CC for the treatment, diagnosis and screening of MOT12-related
CC diseases or conditions characterised by an abnormal interaction
CC between MOT12 and a natural binding partner, especially diabetes,
CC an immune disorder or cancer (claimed).
SQ Sequence 1015 AA;

Query Match 10.6%; Score 99; DB 27; Length 1015;
Best Local Similarity 57.1%; Pred. No. 4.77e+00;
Matches 12; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Db 8 llllllllprvlpapssvp 28
QY 15 LLLLLLLPALLSSGAAAP 35
|||||||:|:|:|:|

RESULT 6
ID W18092 standard; Protein; 1015 AA.
AC W18092;
DT 14-SEP-1997 (first entry)
DE Type I diabetes-associated autoantigen IAR-PTP.
KW Insulin dependent diabetes mellitus-associated autoantigen; IDDM;
KW islet cell antigen-related protein tyrosine phosphatase; IAR-PTP;
KW diagnosis.
OS Homo sapiens.
FH Key Location/Qualifiers
FT peptide 1..21

```

```

FT protein /label= Sig_peptide
FT 22..1015
FT /label= Mat_protein
FT domain 22..614
FT /label= Extracellular_domain
FT domain 615..639
FT /label= Transmembrane_domain
FT domain 640..1015
FT /label= Intracellular_domain
FT active_site 942..954
FT modified_site 564..566
FT /label= Glycosylation
FT /note="potential N-linked glycosylation site"
WO9722694-A2.
PN 26-JUN-1997.
PD 20-DEC-1996; CA0867.
PR 20-DEC-1995; GB-026036.
PR 19-MAR-1996; GB-005710.
PR 27-SEP-1996; GB-020265.
PA (UYST-) UNIV SINGAPORE NAT.
PI Pallen CJ;
PI WPI; 97-341685/31.
DR N-PSDB; T67296.
DR Novel islet cell antigen-related protein tyrosine phosphatase -
PT useful for recognising auto-antibodies associated with insulin
PT dependent diabetes mellitus
PT Claim 4; Page 57-63; 67pp; English.
CC Human islet cell antigen-related protein tyrosine phosphatase
CC (IAR-PTP) (W18092) has the properties of an insulin-dependent
CC diabetes mellitus (IDDM)-associated autoantigen. Its amino acid
CC sequence was deduced from cDNA clones (T67296) obtd. from human
CC pancreas and brain cDNA libraries. Complete, partial (see also
CC W18091) and cytoplasmic region (see also W18093) IAR-PTP
CC polypeptides can be produced using host-vector systems and used,
CC either as an alternative to autoantigen IA-2 or in combination with
CC IA-2, for the diagnosis of IDDM, e.g. to screen for or predict the
CC onset, presence or development of IDDM.
SQ Sequence 1015 AA;

Query Match 10.6%; Score 99; DB 23; Length 1015;
Best Local Similarity 57.1%; Pred. No. 4.77e+00;
Matches 12; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Db 8 llllllllprvlpapssvp 28
QY 15 LLLLLLLPALLSSGAAAP 35
|||||||:|:|:|:|

RESULT 7
ID R85203 standard; Protein; 1337 AA.
AC R85203;
DT 12-FEB-1996 (first entry)
DE huDEP-1.
KW Density enhanced Type III receptor-like protein tyrosine phosphatase;
KW huDEP-1.
OS Homo sapiens.
PN WO9530008-A1.
PD 09-NOV-1995.
PF 03-MAY-1995; U05512.
PR 03-MAY-1994; US-237940.
PA (COLD-) COLD SPRING HARBOR LAB.
PI Oestman A, Tonks NK;
PI WPI; 95-393079/50.
DR N-PSDB; T06027.
DR New density enhanced protein tyrosine phosphatase - used to develop
PT prods. to modify transcription, translation and/or activity of
PT tyrosine phosphatase(s).
PS Claim 4; Page 34-38; 51pp; English.
CC A cDNA clone was obt'd. (see T06027) from a HeLa cell cDNA library
CC that encoded a novel density-enhanced Type III receptor-like PTP,
CC designated huDEP-1 (R85203). huDEP-1 is useful for the study of PTPs
CC and for the development of therapeutic or prophylactic cpds. e.g. for
CC prevention of abnormal or malignant cell growth.

```

```
Query Match          9.8%; Score 91; DB 5; Length 492;
Best Local Similarity 47.6%; Pred. No. 1.78e+01;
Matches 10; Conservative 9; Mismatches 2; Indels 0; Gaps 0;

Db 22 lllllllpslmararpes 42
QY 15 LLLLLLLPALLSGAAQ 35

RESULT 11
ID P94256 standard; protein; 35 AA.
AC P94256;
DT 10-JUN-1997 (revised)
DE Truncated E protein from Phix174.
KW E protein; antimicrobial agent; Phix174.
PN W08900199-A.
PD 12-JAN-1989.
06-JUL-1988; U02265.
06-JUL-1987; US-069653.
(LOU) Louisiana State University Agricultural and Mechanical College.
PI Jaynes JM, Enright FW, White KL;
DR WPI: 89-039653/05.
DR N-PSDB; N91204.
PT Treating or preventing microbial infections - using cecropin(s),
PT attacin(s), lysozyme(s), S protein from lambda phage, E protein
PT from Phix174 or protein from phage 22.
PS Claim 108; page 58; 64pp; English.
CC The truncated E protein is an antimicrobial polypeptide. DNA encoding the
CC protein is used to transform animal cells to express the protein,
CC providing the animals with resistance to microbial infections. The
CC proteins may also be admin. to humans and animals for preventing or
CC treating such infections, eg brucellosis, malarial infection,
CC listeriosis and Chagas' disease.
CC See also P94254-P94258.
CC (Revised entry submitted to correct the format of the patent
CC publication date.)
SQ Sequence 35 AA;

Query Match          9.6%; Score 89; DB 1; Length 35;
Best Local Similarity 55.6%; Pred. No. 2.47e+01;
Matches 10; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 7 wdtlaflllllpsll 24
QY 10 WAAVTLILLLLLPALL 27

JUL 12
ID R67540 standard; Protein; 231 AA.
AC R67540;
DT 05-AUG-1995 (first entry)
DE Mouse flt-3 ligand.
KW Flt-3 ligand; flt3-l; anemia; cancer; AIDS; gene therapy.
FS Mus sp.
FH Key Location/Qualifiers
FT peptide 1..27
FT domain 28..188
FT domain /label- Extracellular_domain
FT domain 189..211
FT domain /label- Transmembrane_domain
FT domain 212..231
FT domain /label- Cytoplasmic_domain
PN EP-627487-A.
PD 07-DEC-1994.
PF 19-MAY-1994; 303575.
PR 24-MAY-1993; US-068394.
PR 12-AUG-1993; US-106463.
PR 25-AUG-1993; US-111758.
PR 03-DEC-1993; US-162407.
PR 07-MAR-1994; US-209502.

Query Match          9.8%; Score 89; DB 13; Length 232;
Best Local Similarity 57.9%; Pred. No. 2.47e+01;
Matches 11; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

Db 7 awspnslllllllspcl 25
QY 9 SWAA-VTLILLLLLPAL 26

Query Match          9.6%; Score 89; DB 12; Length 231;
Best Local Similarity 57.9%; Pred. No. 2.47e+01;
Matches 11; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

Db 7 awspnslllllllspcl 25
QY 9 SWAA-VTLILLLLLPAL 26

RESULT 13
ID R66177 standard; Peptide; 232 AA.
AC R66177;
DT 10-AUG-1995 (first entry)
DE Mouse Mot110/T118 Flt3 ligand peptide fragment.
KW Flt3 ligand; tyrosine kinase receptor ligand.
OS Homo sapiens.
PN W09426891-A.
PD 24-NOV-1994.
PF 18-MAY-1994; U05150.
PR 19-MAY-1993; US-065231.
PR 07-JUL-1993; US-089263.
PR 16-JUL-1993; US-092549.
PR 13-AUG-1993; US-106340.
PR 24-AUG-1993; US-112391.
PR 19-NOV-1993; US-155111.
PR 03-DEC-1993; US-162413.
PA (INEM ) INST NAT SANTE & RECH MEDICALE.
PA (SCHE ) SCHERING CORP.
PI Birnbaum D, Culpepper JA, Hannum CH, Lee FD;
DR N-PSDB; Q79464.
DR WPI: 95-006787/01.
PT New ligand for the Flt3 tyrosine kinase receptor - and related
PT nucleic acid, vectors, host cells and antibodies, useful for
PT treating abnormal cell physiology and drug screening
PT also for diagnosis and drug screening
PS Claim 11; page 79-80; 90pp; English.
CC A cDNA library from the human stromal cell line 29SV48, in
CC pME18S, was screened with an 800 bp fragment derived from
CC mouse clone T118. This fragment encompasses the coding region
CC conserved between two mouse clones, T118 and T110. Approx. 20
CC positive clones were selected and partially sequenced. Two
CC clones, S86 and S109, were found to be approx. 75% homologous
CC to the mouse clones over the first 163 AAs. Clone S86 continued
CC to show homology to T110 until the stop codon, although to a
CC lesser degree, for an overall homology of 66%. Clones T118 and
CC S109 do not show homology to each other or to the other clones
CC after mouse residue 163 (human residue 160). An additional mouse
CC clone designated MB8 has a 29 AA insert at the junction between
CC the common and divergent portions of the mouse ligand.
SQ Sequence 232 AA;

Query Match          9.6%; Score 89; DB 13; Length 232;
Best Local Similarity 57.9%; Pred. No. 2.47e+01;
Matches 11; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

Db 7 awspnslllllllspcl 25
QY 9 SWAA-VTLILLLLLPAL 26
```

WIREH

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MParch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:33:36 1998; Maspar time 3.35 Seconds

128.334 Million cell updates/sec

Output not generated.

Title: >US-08-938-548A-12
Description: (1-28) from US08938548A.pap
Perfect Score: 201
Sequence: 1 RPPGPGQLRQLRLQNGHAGILTM 28

Scoring table: PAM 150

Gap 15

Searched: 124785 seqs, 15338987 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: a-genseq31-2

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27

Statistics: Mean 22.110; Variance 87.356; scale 0.253

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	68	33.8	2192	23	LexA/NumA fusion prot	4.36e+01
2	68	33.8	2272	23	GAL4/HA/NumA fusion p	4.36e+01
3	66	32.8	10	4	"p33" N-terminal (2)	6.36e+01
4	65	32.3	716	18	Retinoid X receptor i	7.68e+01
5	65	32.3	2509	24	Protein (OA-519) cros	7.68e+01
6	64	31.8	256	9	Deduced sequence of h	9.25e+01
7	64	31.8	256	15	MY17 precursor-3.	9.25e+01
8	64	31.8	549	16	KM31-7 precursor.	9.25e+01
9	64	31.8	1477	13	S. cerevisiae scaut2R	9.25e+01
10	64	31.8	1477	22	Saccharomyces cerevis	9.25e+01
11	63	31.3	878	6	DHR23alpha protein.	1.11e+02
12	63	31.3	878	3	Ecdysone receptor.	1.11e+02
13	63	31.3	1841	25	Tyrosine synthase OR	1.11e+02
14	63	31.3	4630	22	Streptomyces venezuel	1.11e+02
15	62	30.8	684	24	Human alpha-1 collage	1.34e+02
16	61	30.3	32	1	New polypeptide based	1.61e+02
17	61	30.3	248	1	Human 32K ASP encoded	1.61e+02
18	61	30.3	248	1	Sequence deduced from	1.61e+02
19	61	30.3	248	3	Genomic sequence of h	1.61e+02

20	61	30.3	248	3	P60441	Plasmid pASPC-SV(10)	1.61e+02
21	61	30.3	248	1	R06331	Human alveolar surfac	1.61e+02
22	61	30.3	248	1	R04215	Human 32K ASP encoded	1.61e+02
23	61	30.3	248	1	P60694	Sequence deduced from	1.61e+02
24	61	30.3	248	2	P70662	35kd pulmonary surfac	1.61e+02
25	61	30.3	248	2	P70663	35kd pulmonary surfac	1.61e+02
26	61	30.3	248	3	P60665	Sequence of human alv	1.61e+02
27	61	30.3	248	3	P60442	Plasmid pASPCq-SV(10)	1.61e+02
28	61	30.3	248	2	R05091	Vector PSP 35K-1A-10	1.61e+02
29	61	30.3	271	1	R04212	Human 32K alveolar su	1.61e+02
30	61	30.3	271	1	P60661	Genomic sequence of h	1.61e+02
31	61	30.3	271	1	R04217	Human 32K ASP encoded	1.61e+02
32	60	29.9	169	8	R43261	Human adipogenesis in	1.93e+02
33	60	29.9	177	8	R43262	Human adipogenesis in	1.93e+02
34	60	29.9	178	18	W02202	Human interleukin-11	1.93e+02
35	60	29.9	178	14	R75337	Human interleukin-11	1.93e+02
36	60	29.9	199	8	R43260	Human adipogenesis in	1.93e+02
37	60	29.9	199	10	R50176	Human interleukin-11	1.93e+02
38	60	29.9	296	5	R26213	Fusion protein of IL-	1.93e+02
39	60	29.9	296	5	R26051	Thioredoxin from PAUT	1.93e+02
40	60	29.9	296	13	R76812	Thioredoxin-IL-11 fus	1.93e+02
41	60	29.9	296	13	R75762	Thioredoxin/Des-Pro-I	1.93e+02
42	60	29.9	405	18	W00084	Human Ras-related pro	1.93e+02
43	60	29.9	510	9	R49835	Thermus aquaticus hea	1.93e+02
44	60	29.9	530	14	R76480	Virulence-associated	1.93e+02
45	60	29.9	1052	15	R88576	Human focal adhesion	1.93e+02

ALIGNMENTS

RESULT 1
ID W21732 standard; Protein; 2192 AA.
AC W21732;
DT 01-OCT-1997 (first entry)
DE LexA/NumA fusion protein.
KW NIP-1; NIP-2; NumA; nuclear mitotic apparatus; NumA interacting protein;
KW cell division; proliferation; antibody; Ab; detection;
OS Homo sapiens.
FH Key
FT domain 1..87 Location/Qualifiers
FT /label= LexA_DNA_binding_domain
FT peptide 88..94
FT /label= Polylinker
FT protein 95..2192
FT /label= Residues_18-2116_of_NumA
FT region 285..1784
FT /label= Coiled_coil_region
PN W09640917-A1.
PD 19-DEC-1996.
PP 07-JUN-1996; U09504.
PP 07-JUN-1995; US-478408.
PA (UYVA) UNIV YALE.
PI McPherson SMG, Snyder MP;
WI 97-077270/07.
DR N-PSDB; T77783.
DR N-PSDB; T77783.
PT New nucleic acid encoding nuclear mitotic appts. interacting
PT proteins - useful for modulating cell division and proliferation and
PT in diagnosis
PS Claim 15; Page 42-50; 78pp; English.
CC The sequences given in W21731-32 represent fusion proteins which contain
CC NumA (nuclear mitotic apparatus). The fusion proteins were used in
CC the identification of NumA interacting proteins (NIP's) (see also
CC W21729-30). Compounds which interfere with the interaction of NumA
CC with a known NIP are used to modulate cell division and/or proliferation.
CC Ab, raised conventionally using NIP-1 or -2 as immunogen, are used to
CC detect NIP (or their complexes) and to block their activity for
CC diagnostic or therapeutic use, e.g. to detect defective NumA or NIP
CC which may be markers for aberrant (including malignant) cell growth
CC (which can also be detected by nucleic acid sequencing). Also where
CC malignancy is related to defects in NumA or NIP, it can be treated by
CC administration of the appropriate functional protein.
SQ Sequence 2192 AA;

DT 16-JAN-1998 (first entry)
 DE Protein (OA-519) cross-reactive with hpr gene product.
 KW OA-519; cross-reaction; haptoglobin related; hpr; antibody;
 KW epitope; haptoglobin 1; haptoglobin 2; cancer; breast cancer;
 KW prognosis assay.
 OS Homo sapiens.
 PN US565874-A.
 PD 09-SEP-1997.
 PF 17-JAN-1989; 297722.
 PR 24-JAN-1994; US-188426.
 PR 17-JAN-1989; US-297722.
 PR 04-DEC-1990; US-622407.
 PR 26-JUL-1991; US-735522.
 PR 24-JUL-1992; US-917716.
 PR 26-JUL-1993; US-096908.
 PR 05-JUN-1995; US-469005.
 PA (UYJO) UNIV JOHNS HOPKINS.
 PI Kuhlida FP, Pasternack GR;
 DR WPI: 97-469516/43.
 DR N-PSDB; T88206.
 DN DNA encoding protein cross-reactive with hpr gene product - useful
 to raise antibodies reactive with epitope(s) found on hpr gene
 product, useful in cancer, especially breast cancer, prognosis
 assays
 PS Claim 2: Columns 43-60; 68pp; English.#
 CC The present sequence is a protein (OA-519) cross-reactive with
 the haptoglobin related (hpr) gene product. OA-519 can be used to
 raise antibodies reactive with epitopes found on the hpr gene
 product, but not on haptoglobin 1 or 2, useful in cancer,
 CC especially breast cancer, prognosis assays.
 CC Sequence 2509 AA;
 SQ

Query Match 32.3%; Score 65; DB 24; Length 2509;
 Best Local Similarity 29.6%; Pred. No. 7.69e+01;
 Matches 8; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

Db 411 ppapaphatiprllrasgrtpeavqkl 437
 QY 2 PGPPGLQRLRLQANGNHAAGILTM 28

RESULT 6
 ID R45403 standard; Protein; 256 AA.
 AC R45403;
 DT 18-JUL-1994 (first entry)
 DE Deduced sequence of human proteinase-3 (PR-3)
 KW TNF convertase; proteinase-3; PR-3; tumour necrosis factor;
 KW proTNF.
 OS Homo sapiens.
 PN WO9400555-A.
 PD 06-JAN-1994.
 PF 25-JUN-1993; U06120.
 PR 25-JUN-1992; US-905546.
 PA (CETU) CETUS ONCOLOGY CORP.
 PI Halenbeck RF, Jewell DA, Kothe KE, Kriegler M, Perez C;
 DR WPI: 94-026195/03.
 DR N-PSDB; Q54498.
 PT Cpd, which inhibit formation of mature tumour necrosis factor
 from its precursor - identified using TNF convertase, e.g.
 PT mutin(s), antibodies or peptide phosphonate(s), for preventing
 and treating sepsis, AIDS, auto-immune disease etc.
 PS Disclosure; Fig 2; 69pp; English.
 CC ProTNF refers to TNF having a molecular weight of about 26,000,
 CC which is the proform of TNF. ProTNF is cleaved to a lower
 CC molecular weight 'mature' form, prof. 17kd, which, in its multimeric
 CC (usually trimeric) form, is substantially involved in producing life-
 CC threatening physiological changes associated with sepsis. ProTNF is
 CC cleaved by convertase. One TNF convertase is serine protease
 CC proteinase-3, also called PR-3, P-29B or myeloblastin. A suitable
 CC source of convertase is the HL60 cell line (or extracts, or the
 CC culture media in which it is grown). The convertase produced by
 CC HL60 has been sequenced and is identical to the known lymphocyte
 CC serine protease PR-3 which has other activities unrelated to TNF

CC processing.
 SQ Sequence 256 AA;
 Query Match 31.8%; Score 64; DB 9; Length 256;
 Best Local Similarity 38.5%; Pred. No. 9.25e+01;
 Matches 10; Conservative 7; Mismatches 9; Indels 0; Gaps 0;
 Db 4 rpspalasvllallsgaaraaiv 29
 QY 1 RPGPGLQRLRLQANGNHAAGIL 26
 RESULT 7
 ID R85639 standard; Protein; 256 AA.
 AC R85639;
 DT 23-APR-1996 (first entry)
 DE MY17 preproPR-3.
 KW PR-3; preproPR-3; MY17; human neutrophil protease-3; serine protease;
 KW tumour necrosis factor alpha; TNFalpha; HL60; MY17; B cell; T cell;
 KW tumour necrosis factor alpha convertase; cytokine; septic shock;
 KW rheumatoid arthritis; cachexia; cerebral malaria; graft-host disease;
 KW ischaemia/reperfusion injury; autoimmune disease; AIDS.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT peptide 1..25
 FT /note= "leader sequence present only in preproPR-3"
 FT peptide 26..27
 FT /note= "dipeptide present in proPR-3"
 FT protein 28..256
 FT /note= "mature PR-3"
 PN WO9524501-A1.
 PD 14-SEP-1995.
 PF 02-MAR-1995; U02513.
 PR 07-MAR-1994; US-208574.
 PR 19-APR-1994; US-230428.
 PR 27-FEB-1995; US-394600.
 PR 28-FEB-1995; US-395456.
 PA (CETU) CETUS ONCOLOGY CORP.
 PI Halenbeck RF, Jewell DA, Kothe KE, Kriegler M, Perez C;
 DR WPI: 95-328287/42.
 DR N-PSDB; T02565.
 PT Identification of inhibitors of mature TNFalpha prodn. - useful for
 treatment of septic shock, rheumatoid arthritis, etc..
 PS Example 2; Page 82; 96pp; English.
 CC This sequence represents the preproPR-3. PR-3 is active recombinant
 CC human neutrophil protease-3. PR-3 is a serine protease, and is a tumour
 CC necrosis factor alpha (TNFalpha) convertase. The cDNA encoding this
 CC sequence was isolated from the HL60 cell clone MY17. The mature PR-3 can
 CC be used in the method of the invention for identifying agents that
 CC inhibit cleavage of proTNFalpha to mature TNFalpha. In the method,
 CC proTNFalpha is incubated with PR-3 (or another TNFalpha convertase), and
 CC the cleavage of the proTNFalpha is measured by a colourimetric assay.
 CC This is then repeated in the presence of a test compound that is thought
 CC to inhibit this process. The results of the two reactions are then
 CC compared to determine whether the test compound is an inhibitor. The
 CC cleavage inhibitors can be used in the treatment of septic shock,
 CC rheumatoid arthritis, cachexia, cerebral malaria, ischaemia/reperfusion
 CC injury, graft-host disease, autoimmune diseases, and AIDS. PR-3 can be
 CC used to treat unwanted B cell/T cell interaction by contacting it with
 CC T cells to cause the release of membrane-bound cytokines.
 SQ Sequence 256 AA;

Query Match 31.8%; Score 64; DB 15; Length 256;
 Best Local Similarity 38.5%; Pred. No. 9.25e+01;
 Matches 10; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Db 4 rpspalasvllallsgaaraaiv 29
 QY 1 RPGPGLQRLRLQANGNHAAGIL 26
 RESULT 8
 ID R92050 standard; Protein; 549 AA.


```

FT      /label= GXGX'Y'_motif
FT      /note= "Claim 1"
FT peptide
FT      329..334
FT      /label= GXGX'Y'_motif
FT      /note= "Claim 1"
FT peptide
FT      335..340
FT      /label= GXGX'Y'_motif
FT      /note= "Claim 1"
FT peptide
FT      354..359
FT      /label= GXGX'Y'_motif
FT      /note= "Claim 1"
FT peptide
FT      360..365
FT      /label= GXGX'Y'_motif
FT      /note= "Claim 1"
FT peptide
FT      366..372
FT      /label= GXGX'Y'_motif
FT      /note= "Claim 1"
FT peptide
FT      523..528
FT      /label= GXGX'Y'_motif
FT      /note= "Claim 1"
FT peptide
FT      542..547
FT      /label= GXGX'Y'_motif
FT      /note= "Claim 1"
FT peptide
FT      590..595
FT      /label= GXGX'Y'_motif
FT      /note= "Claim 1"
FT
FT      US5643783-A.
FT      01-JUL-1997.
FT      01-DEC-1993; 159784.
FT      01-DEC-1993; US-159784.
FT      (HARD ) HARVARD COLLEGE.
FT      Oh SP, Olsen BR:
FT      WPI: 97-350247/32.
FT      N-PSDB; T84484.
FT      Nucleic acid encoding human alpha-1 collagen - for production of
FT      recombinant alpha-1 collagen, for use in the treatment of cartilage
FT      degeneration
FT      Claim 1; Column 23-30; 35pp; English.
FT      Novel human type alpha-1 (XVIII) collagen is characterised by
FT      10 triple helical domains containing the GXGX'Y' motif (where X,
FT      Y, X' and Y' represent any amino acid), the helical domains being
FT      separated and flanked by non-triple helical regions which may
FT      provide flexibility. Alpha-1 collagen is expressed in multiple
FT      tissues, especially liver, lung and kidney. A claimed plasmid
FT      comprising alpha-1 collagen nucleic acid (see T84484) and an
FT      expression control sequence can be used to express recombinant
FT      cells. The alpha-1 collagen may be used to treat a patient
FT      suffering from a disease associated with cartilage degradation, and
FT      for supplementing collagen. It can also be used as a connective
FT      tissue filler (e.g. for plastic surgery), can be interposed between
FT      a dermal equivalent and skin to improve adhesion, and as a
FT      substrate on which to grow epithelium. The addition of
FT      glycosaminoglycans makes alpha-1 collagen a better carrier for
FT      fibroblast growth factor, and imparts greater tensile strength.
FT      Sequence 684 AA;

```

```

Query Match      30.8%; Score 62; DB 24; Length 684;
Best Local Similarity 88.9%; Pred. No. 1.34e+02;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Db 213 rpgpglpg 221
   |||||
QY 1 RFPFPLQG 9

```

Search completed: Thu Jul 30 09:33:56 1998
Job time : 20 secs.

(TM)

```

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
on: Thu Jul 30 09:36:05 1998; MasPar time 1.18 Seconds
138.599 Million cell updates/sec
Molecular output not generated.

```

Ult No.	Score	Query		Length	DB	ID	Description	Pred. No.
		Match	Score					
1	66	32.8	10	2	PCT-US91-0	Sequence 3, Applicatio	2.07e+01	
2	65	32.3	716	2	PCT-US95-1	Sequence 4, Applicatio	2.50e+01	
3	65	32.3	2509	1	US-08-459-	Sequence 10, Applicati	2.50e+01	
4	62	30.8	1694	1	US-08-494-	Sequence 2, Applicatio	4.40e+01	
5	60	29.9	199	1	US-08-115-	Sequence 2, Applicatio	6.39e+01	
6	60	29.9	199	1	US-08-792-	Sequence 8, Applicatio	6.39e+01	
7	60	29.9	199	1	US-07-941-	Sequence 2, Applicatio	6.39e+01	
8	60	29.9	199	2	PCT-US93-0	Sequence 4, Applicatio	6.39e+01	
9	60	29.9	199	1	US-07-949-	Sequence 2, Applicatio	6.39e+01	
10	60	29.9	199	1	US-08-017-	Sequence 4, Applicatio	6.39e+01	
11	60	29.9	296	1	US-07-745-	Sequence 14, Applicati	6.39e+01	
12	60	29.9	296	1	US-08-115-	Sequence 4, Applicatio	6.39e+01	
13	60	29.9	296	1	US-08-165-	Sequence 14, Applicati	6.39e+01	
14	60	29.9	296	1	US-07-921-	Sequence 14, Applicati	6.39e+01	
15	60	29.9	296	1	US-07-941-	Sequence 4, Applicatio	6.39e+01	
16	60	29.9	296	2	PCT-US93-0	Sequence 4, Applicatio	6.39e+01	
17	60	29.9	296	2	PCT-US94-1	Sequence 14, Applicati	6.39e+01	
18	60	29.9	405	2	US-07-688-	Sequence 14, Applicati	6.39e+01	
19	60	29.9	405	2	PCT-US91-0	Sequence 14, Applicati	6.39e+01	
20	59	29.4	176	2	PCT-US96-1	Sequence 6, Applicatio	7.69e+01	
21	58	28.9	26	1	US-07-776-	Sequence 25, Applicati	9.12e+01	
22	58	28.9	27	1	US-07-924-	Sequence 10, Applicati	9.12e+01	
23	58	28.9	27	1	US-08-063-	Sequence 43, Applicati	9.12e+01	

RESULT 9
ID US-07-949-516A-4 STANDARD; PRT; 199 AA.
XX AC xxxxxx
XX DT
XX DE
XX DE Sequence 4, Application US/07949516A
XX DE Sequence 4, Application US/07949516A
CC CC Patent No. 5700664
CC CC GENERAL INFORMATION:
CC CC APPLICANT: Yang, Yu-Chung
CC CC APPLICANT: Bennett, Frances
CC CC TITLE OF INVENTION: A MAMMALIAN CYTOKINE, IL-11
CC CC NUMBER OF SEQUENCES: 4
CC CC CORRESPONDENCE ADDRESS:
CC CC ADDRESSEE: LEGAL AFFAIRS
CC CC STREET: 87 Cambridgepark Drive
CC CC CITY: Cambridge
CC CC STATE: MA
CC CC COUNTRY: USA
CC CC ZIP: 02140
CC CC COMPUTER READABLE FORM:
CC CC MEDIUM TYPE: Floppy disk
CC CC COMPUTER: IBM PC compatible
CC CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CC CURRENT APPLICATION DATA:
CC CC APPLICATION NUMBER: US/07/949,516A
CC CC FILING DATE: 19-NOV-1992
CC CC CLASSIFICATION: 530
CC CC ATTORNEY/AGENT INFORMATION:
CC CC NAME: Meindert, M. C.
CC CC REGISTRATION NUMBER: 31,544
CC CC REFERENCE/DOCKET NUMBER: 5174BPCT
CC CC INFORMATION FOR SEQ ID NO: 4:
CC CC SEQUENCE CHARACTERISTICS:
CC CC LENGTH: 199 amino acids
CC CC TYPE: amino acid
CC CC TOPOLOGY: linear
CC CC MOLECULE TYPE: protein
CC CC SEQUENCE 199 AA; 21429 MW; 188641 CN;
SQ
Query Match 29.9%; Score 60; DB 1; Length 199;
Best Local Similarity 70.0%; Pred. No. 6.39e+01;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
129 LQARLDRLRL 138
II:II:II:
7 LQGRLOQLLQ 16
OY

RESULT 10
ID US-08-017-522A-4 STANDARD; PRT; 199 AA.
XX AC xxxxxx
XX DT
XX DE
XX DE Sequence 4, Application US/08017522A
XX DE Patent No. 5371193
CC CC GENERAL INFORMATION:
CC CC APPLICANT: BENNETT, FRANCES K
CC CC APPLICANT: PAUL, STEPHAN R
CC CC APPLICANT: YANG, YU-CHUNG
CC CC TITLE OF INVENTION: A MAMMALIAN CYTOKINE, IL-11
CC CC NUMBER OF SEQUENCES: 4
CC CC CORRESPONDENCE ADDRESS:
CC CC ADDRESSEE: LEGAL AFFAIRS - GENETICS INSTITUTE, INC.
CC CC STREET: 87 CAMBRIDGEPARK DRIVE

CC CITY: CAMBRIDGE
CC STATE: MA
CC COUNTRY: US
CC ZIP: 02140
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/017,522A
CC FILING DATE: 19930212
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: MEINERT, M C
CC REGISTRATION NUMBER: 31,544
CC REFERENCE/DOCKET NUMBER: GI 5174A-DIV
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 876-1210 X8574
CC TELEFAX: (617) 876-5851
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 199 amino acids
CC TYPE: AMINO ACID
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 199 AA; 21429 MW; 188641 CN;
SQ
Query Match 29.9%; Score 60; DB 1; Length 199;
Best Local Similarity 70.0%; Pred. No. 6.39e+01;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
129 LQARLDRLRL 138
II:II:II:
7 LQGRLOQLLQ 16
OY

RESULT 11
ID US-07-745-382-14 STANDARD; PRT; 296 AA.
XX AC xxxxxx
XX DT
XX DE
XX DE Sequence 14, Application US/07745382
XX DE Patent No. 5270181
CC CC GENERAL INFORMATION:
CC CC APPLICANT: McCoy, John
CC CC APPLICANT: Lavallie, Edward
CC CC TITLE OF INVENTION: Peptide and Protein Fusions To
CC CC TITLE OF INVENTION: Thioredoxin and Thioredoxin-Like Molecules
CC CC NUMBER OF SEQUENCES: 24
CC CC CORRESPONDENCE ADDRESS:
CC CC ADDRESSEE: Genetics Institute, Inc.
CC CC STREET: 87 Cambridgepark Drive
CC CC CITY: Cambridge
CC CC STATE: Massachusetts
CC CC COUNTRY: U.S.A.
CC CC ZIP: 02140
CC CC COMPUTER READABLE FORM:
CC CC MEDIUM TYPE: Floppy disk
CC CC COMPUTER: IBM PC compatible
CC CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CC CURRENT APPLICATION DATA:
CC CC APPLICATION NUMBER: US/07/745,382
CC CC FILING DATE: 19910814
CC CC CLASSIFICATION: 435
CC CC PRIOR APPLICATION DATA:
CC CC APPLICATION NUMBER: US 07/652,531

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:34:14 1998; MasPar time 4.39 Seconds
233.145 Million cell upd

ular output not generated.

```
>US-08-938-548A-12
Description:
Perfect Score: 201
Sequence: 1 RGPGGLOGRLQRLLOANGNHAAGILTM 28
```

Scoring table: PAM 150

[illegible]

Searched: 120441 seqs, 36531193 residues

Post-processor: Minimum $M_2 + \alpha b$ Post-processing: Minimum Match 0%
Listing first 45 summaries

TEST

Database: pir56

1:pir1 2:pir2 3:pir3 4:pir4 5:nrl3d

Statistics: Mean 30.142; Variance 55.308; scale 0.545

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description	Pred. No.
1	71	35.3	187	2 A41654	superoxide dismutase	2.25e+00
2	71	35.3	187	2 A41654	superoxide dismutase	2.25e+00
3	68	33.8	736	2 S47044	MALDP protein - mouse	6.08e+00
4	68	33.8	879	2 H64898	hypothetical protein	6.08e+00
5	68	33.8	2115	2 S23647	NUMA protein - human	6.08e+00
6	67	33.3	178	2 E69913	hypothetical protein	8.42e+00
7	67	33.3	442	2 JH0496	transcription initiat	8.42e+00
8	67	33.3	442	2 S11712	transcription initiat	8.42e+00
9	67	33.3	510	2 S41307	transcription initiat	8.42e+00
10	67	33.3	525	2 JN0443	transcription initiat	8.42e+00
11	67	33.3	528	2 JN0445	transcription initiat	8.42e+00
12	67	33.3	1157	2 A55132	PAS1 protein - yeast	8.42e+00
13	66	32.8	150	2 D69081	deoxyuridine 5-tripho	1.16e+01
14	66	32.8	253	2 S49183	hypothetical protein	1.16e+01
15	66	32.8	319	1 S19248	RNA-directed DNA poly	1.16e+01
16	65	32.3	1692	2 A33988	adenylate cyclase [EC	1.60e+01
17	65	32.3	2509	2 G01880	fatty-acid synthase (1.60e+01
18	64	31.8	239	2 S25618	hypothetical protein	2.20e+01
19	64	31.8	256	1 PRH3	proteinase 3 (EC 3.4.	2.20e+01
20	64	31.8	263	2 S51155	hypothetical protein	2.20e+01
21	64	31.8	439	2 S61858	hrpE protein - pseudo	2.20e+01
22	64	31.8	535	2 A46101	protein-tyrosine-phos	2.20e+01
23	64	31.8	548	2 B46101	protein-tyrosine-phos	2.20e+01

```

QY      7 LQGRLLQLQANGNHAAGILTM 28

RESULT      8
ENTRY       S11712      #type complete
TITLE       transcription initiation factor sigma hrdB - Streptomyces coelicolor
ORGANISM    #formal_name Streptomyces coelicolor
DATE        03-May-1994 #sequence_revision 20-Feb-1995 #text_change 08-Sep-1997
ACCESSIONS  S11712
REFERENCE   S11711
            Tanaka, K.; Shina, T.; Takahashi, H.
            #submission Submitted to the EMBL Data Library, May 1990
            #description Multiple homolog genes for principal sigma subunit of Streptomyces coelicolor A3(2).
            #accession S11712
            #molecule_type DNA
            ##residues 1-442 #label EMB
            ##cross-references EMBL:X52983; NID:g48744; PID:g48745
            ##experimental_source strain A3(2)
GENETICS
#gene       hrdB
CLASSIFICATION #superfamily transcription initiation factor sigma katF;
               transcription initiation factor sigma katF homology
KEYWORDS     DNA binding; sigma factor; transcription initiation
FEATURE      211-437
            #domain transcription initiation factor sigma katF
            homology #label kTF
SUMMARY      #length 442 #molecular-weight 48413 #checksum 4794

Query Match          33.3%; Score 67; DB 2; Length 442;
Best Local Similarity 31.8%; Pred. No. 8.42e+00;
Matches              7; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

Db      370 LQGLHSVLDTLSERDGVVSM 391
|| :|: |: :: |:::||
QY      7 LQGRLLQLQANGNHAAGILTM 28


RESULT      9
ENTRY       S41307      #type complete
TITLE       transcription initiation factor sigma - Streptomyces griseus
ORGANISM    #formal_name Streptomyces griseus
DATE        06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Sep-1997
ACCESSIONS  S41307
REFERENCE   S41306
            Marcos, A.T.; Diez, B.; Gutierrez, S.; Fernandez, F.J.; Velasco, J.; Martin, J.F.
            #submission Submitted to the EMBL Data Library, December 1993
            #description Organization and expression of the hrdB-sprC gene cluster of streptomyces griseus encoding a sigma factor protein and a serine protease. Role on growth and sporulation of streptomyces.
            #accession S41307
            ##status preliminary
            ##molecule_type DNA
            ##residues 1-510 #label MAR
            ##cross-references EMBL:X75952; NID:g440164; PID:g581664
GENETICS
#start_codon GTG
CLASSIFICATION #superfamily transcription initiation factor sigma katF;
               transcription initiation factor sigma katF homology
KEYWORDS     DNA binding; sigma factor; transcription initiation
FEATURE      279-505
            #domain transcription initiation factor sigma katF
            homology #label kTF
SUMMARY      #length 510 #molecular-weight 55795 #checksum 1415

Query Match          33.3%; Score 67; DB 2; Length 510;
Best Local Similarity 31.8%; Pred. No. 8.42e+00;
Matches              7; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

```

Search completed: Thu Jul 30 09:34:34 1998
Job time : 20 secs.

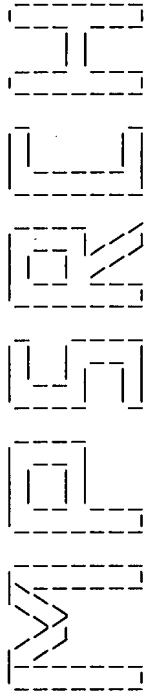
```
SUMMARY
#length 253 #molecular-weight 27305 #checksum 1526
```

```

RESULT 15
ENTRY
TITLE
    ERNATE_NAMES
    DNA nucleotidyltransferase (RNA-directed); reverse
    transcriptase; revertase
    #formal_name Escherichia coli retron Ec107
    #formal_name Escherichia coli retron Ec107
    20-Feb-1995 #sequence_revision 15-Oct-1996 #text_change
    05-Sep-1997
    SI9248
    SI9248
    SI9248
    SI9248
    SI9248
    SI9248
    #authors
    #journal
    #volume
    #issue
    #title
    #cross-references
    Herzer, P.J.; Inouye, S.; Inouye, M.
    Mol. Microbiol. (1992) 6:345-354
    Retron Ec107 is inserted into the Escherichia coli genome by
    replacing a palindromic 34bp intergenic sequence.
    #cross-references
    UID:92204001

```

Query Match 32.8%; Score 66; DB 1; Length 319;



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:34:53 1998; MasPar time 2.66 Seconds
263.713 Million cell updates/sec
Similar output not generated.

Title: >US-08-938-548A-12
Description: (1-28) from US08938548A.pep
Perfect Score: 201
Sequence: 1 RGPFGQLRQLRLQANGNHAAGILTM 28

Scoring table: PAM 150
Gap 15

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 31.553; Variance 48.558; scale 0.650

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	71	35.3	187	1	SODC_HAEP A SUPEROXIDE DISMUTASE P	5.60e-01
2	71	35.3	187	1	SODC_HAEP A SUPEROXIDE DISMUTASE L	5.60e-01
3	69	34.3	1992	1	TR12_HUMAN THYROID RECEPTOR INTER	1.20e+00
4	68	33.8	736	1	ALD_MOUSE ADRENOLEUKODYSTROPHY P	1.74e+00
5	68	33.8	879	1	YDBH_ECOLI HYPOTHETICAL 96.8 KD P	1.74e+00
6	67	33.3	79	1	NIFU_FRAAL NIFU PROTEIN.	2.53e+00
7	67	33.3	442	1	HRDE_STRCO RNA POLYMERASE PRINCIP	2.53e+00
8	67	33.3	1157	1	PEX1_PICPA PEROXISOME BIOSYNTHESI	2.53e+00
9	66	32.8	398	1	VE2_HPV63 REGULATORY PROTEIN E2.	3.65e+00
10	65	32.3	307	1	CC35_CAEEL CUTICLE COLLAGEN 36.	5.25e+00
11	65	32.3	368	1	COAL_POVLY COAT PROTEIN VPI.	5.25e+00
12	65	32.3	799	1	AFSK_STRCO SERINE/THREONINE PROTE	5.25e+00
13	65	32.3	1692	1	CYRA_SCPO ADENYLATE CYCLASE (EC	5.25e+00
14	64	31.8	199	1	IL11_MOUSE INTERLEUKIN-11 PRECURS	7.52e+00
15	64	31.8	256	1	PRN3_HUMAN MYELOBLASTIN PRECURSOR	7.52e+00
16	64	31.8	624	1	SIR_SYN7 SULFITE REDUCTASE (FER	7.52e+00
17	64	31.8	1477	1	FORL_YEAST OLIGOMYCIN RESISTANCE	7.52e+00
18	63	31.3	394	1	YPE1_RHROU HYPOTHETICAL 28 KD PRO	1.07e+01
19	63	31.3	255	1	RT04_YEAST MITOCHONDRIAL 40S RIBO	1.07e+01
20	63	31.3	702	1	YCBY_ECOLI HYPOTHETICAL 78.9 KD P	1.07e+01
21	63	31.3	878	1	ECR_DRONE ECDYSONE RECEPTOR (ECD	1.07e+01
22	62	30.8	256	1	YREC_SYN2 HYPOTHETICAL 28.7 KD P	1.52e+01
23	62	30.8	380	1	F812_MOUSE FACTOR VIII INTRON 22	1.52e+01

Query Match 35.3%; Score 71; DB 1; Length 187;
Best Local Similarity 35.7%; Pred. No. 5.60e-01;

RESULT ID	SODC_HAEP A	STANDARD;	PRT;	187 AA.
AC	P25842;			
DT	01-MAY-1992 (REL. 22, CREATED)			
DT	01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)			
DT	01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)			
DE	SUPEROXIDE DISMUTASE PRECURSOR (CU-ZN) (EC 1.15.1.1).			
GN	SODC.			
OS	HAEMOPHILUS PARAINFLUENZAE.			
OC	PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;			
OC	PASTEURELLACEAE.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-1391;			
RX	MEDLINE; 92041655.			
RA	KROLL J.S., LANGFORD P.R., LOYNDS B.M.;			
RL	J. BACTERIOL. 173:7449-7457(1991).			
CC	- FUNCTION: DESTROYS RADICALS WHICH ARE NORMALLY PRODUCED WITHIN THE CELLS AND ARE TOXIC TO BIOLOGICAL SYSTEMS.			
CC	- FUNCTION: MAY CONFER SURVIVAL ADVANTAGE BY ACCELERATING DISMUTATION OF SUPEROXIDE OF ENVIRONMENTAL ORIGIN TO HYDROGEN PEROXIDE, DISRUPTIVE TO THE NORMAL MUOCILIARY CLEARANCE PROCESS IN THE HOST.			
CC	- CATALYTIC ACTIVITY: 2 PEROXIDE RADICAL + 2 H(+) - O(2) + H(2)O(2).			
CC	- SUBUNIT: HOMODIMER.			
CC	- SUBCELLULAR LOCATION: PERIPLASMIC (POTENTIAL).			
CC	- SIMILARITY: BELONGS TO THE CU-ZN SUPEROXIDE DISMUTASE FAMILY.			
DR	EMBL; M84013; G148884;			
DR	PIR; B41654; B41654.			
DR	HSSP; P00445; LSDY.			
DR	PROSITE; PS00087; SOD_CU_ZN_1; 1.			
DR	PROSITE; PS00332; SOD_CU_ZN_2; 1.			
KW	OXIDOREDUCTASE; COPPER; ZINC; PERIPLASMIC; SIGNAL.			
FT	SIGNAL 1 23			
FT	CHAIN 24 187			
FT	METAL 80 80			
FT	METAL 82 82			
FT	METAL 105 105			
FT	METAL 114 114			
FT	METAL 123 123			
FT	METAL 126 126			
FT	METAL 161 161			
FT	DISULFID 87 183			
FT	SEQUENCE 187 AA; 3DDD4297 CRC32;			

ALIGNMENTS

SQ SEQUENCE 736 AA; 81858 MW; D373B00E CRC32;

Query Match 33.8%; Score 68; DB 1; Length 736;
Best Local Similarity 44.4%; Pred. No. 1.74e+00;
Matches 12; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Db 704 AGIPKMGRLQRLQLGAAAPVQPL 730

QY 2 PGPPGLQRLQLQANGNHAAGILTM 28

RESULT 5

ID YDBH_ECOLI STANDARD; PRT; 879 AA.

AC P52645; P77502; P76855;

DT 01-OCT-1996 (REL. 34, CREATED)

DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL 96.8 KD PROTEIN IN LDHA-TYNA INTERGENIC REGION.

GN YDBH

OS ESCHERICHIA COLI.

PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
ENTEROBACTERIACEAE.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;

RA BLATNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;

RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-K12;

RA AIBA H., BABA T., FUJITA K., HAYASHI K., HONJO A., HORIUCHI T.,

RA IKEMOTO K., INADA T., ISONO K., ITOH T., KANAI K., KASAI H.,

RA KASHIMOTO K., KIM S., KIMURA S., KITAGAWA M., KITAKAWA M., MAKINO K.,

RA MASUDA S., MIKI T., MIZOBUCHI K., MORI H., MOTOMURA K., NAKAMURA Y.,

RA NASHIMOTO H., NISHIO Y., OSHIMA T., SAITO N., SAMPEI G., SEKI Y.,

RA TAGAMI H., TAKEMOTO K., WADA C., YAMAMOTO Y., YANO M.,

RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.

RN [3]

RP SEQUENCE OF 1-331 FROM N.A.

RC STRAIN-K12;

RA BUNCH P.K., MAT-JAN F., LEE N.A., DEAYALA B.A., CLARK D.P.;

RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.

RN [4]

RP IDENTIFICATION.

RA RUDD K.E.;

RL UNPUBLISHED OBSERVATIONS (MAR-1996).

DR EMBL; A8000235; G1787646; -

DR EMBL; D90776; G1742260; -

DR EMBL; D90777; G1742264; -

EMBL; U36928; -; NOT_ANNOTATED_CDS.

ECOGENE; EGI3180; YDBH.

KW HYPOTHETICAL PROTEIN.

FT CONFLICT 36 36 I -> L (IN REF. 3).

SQ SEQUENCE 879 AA; 96834 MW; 8BFD7CF3 CRC32;

Query Match 33.8%; Score 68; DB 1; Length 879;

Best Local Similarity 45.8%; Pred. No. 1.74e+00;

Matches 11; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

Db 396 GVDGRQLAQHLENELGDFVLHM 419

QY 6 GLQGRQLRLQANGNHAAG-ILTM 28

RESULT 6

ID NIFU_FRAAL STANDARD; PRT; 79 AA.

AC P46045;

DT 01-NOV-1995 (REL. 32, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DE NIFU PROTEIN.

GN NIFU.

OS FRANKIA ALNI.

OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; FRANKIACEAE.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-CFII;

RX MEDLINE; 95369734.

RA HARRIOTT O.T., HOSTED T.J., BENSON D.R.;

RL GENE 161:63-67(1995).

DR EMBL; L29299; G497438; -

KW NITROGEN FIXATION.

SQ SEQUENCE 79 AA; 8624 MW; 4CF1624F CRC32;

Query Match 33.3%; Score 67; DB 1; Length 79;

Best Local Similarity 56.3%; Pred. No. 2.53e+00;

Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 29 RRGAPADPRRLRLQ 44

QY 1 RGPFGGLQGRQLRLQ 16

RESULT 7

ID HRDB_STRCO STANDARD; PRT; 442 AA.

AC P18183;

DT 01-NOV-1990 (REL. 16, CREATED)

DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)

DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)

DE RNA POLYMERASE PRINCIPAL SIGMA FACTOR HRDB.

GN HRDB.

OS STREPTOMYCES COELICOLOR.

OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; STREPTOMYCETACEAE.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-A3(2);

RX MEDLINE; 89058616.

RA TANAKA K., SHIINA T., TAKAHASHI H.;

RL SCIENCE 242:1040-1042(1988).

CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES

ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND

THEN IS RELEASED.

CC -1- SIMILARITY: TO OTHER SIGMA FACTORS THAT DO NOT BELONG TO THE

SIGMA-54 FAMILY.

EMBL; X52983; G48745; -

PIR; S11712; S11712.

DR PROSITE; PS00715; SIGMA70_1; 1.

DR PROSITE; PS00716; SIGMA70_2; 1.

KW TRANSCRIPTION REGULATION; SIGMA FACTOR; DNA-DIRECTED RNA POLYMERASE;

FT DNA_BIND 233 246 POLYMERASE CORE BINDING (POTENTIAL).

FT DNA_BIND 403 422 H-T-H MOTIF (BY SIMILARITY).

SQ SEQUENCE 442 AA; 48413 MW; 4720321F CRC32;

Query Match 33.3%; Score 67; DB 1; Length 442;

Best Local Similarity 31.8%; Pred. No. 2.53e+00;

Matches 7; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

Db 370 LQEQHSHVLDLTLSEAGVVM 391

QY 7 LQGRQLRLQANGNHAAGILTM 28

RESULT 8

ID PEX1_PICPA STANDARD; PRT; 1157 AA.

AC P46463;

DT 01-NOV-1995 (REL. 32, CREATED)

DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)

DE PEROXISOME BIOSYNTHESIS PROTEIN PASI (PEROXIN-1).

GN PEX1 OR PASI.

OS PICHIA PASTORIS (YEAST).

OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-21-1;

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE SERINE/THREONINE PROTEIN KINASE AFSK (EC 2.7.1.-).
 GN AFSK.
 OS STREPTOMYCETES COELICOLOR.
 CC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; STREPTOMYCETACEAE.
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RX MEDLINE: 94341568.
 RA MATSUMOTO A., HONG S.K., ISHIZUKA H., HORINOCHI S., BEPPU T.;
 RL GENE 146:47-56(1994).
 RN [2]
 RC SEQUENCE FROM N.A.
 RA UMEYAMA T., HORINOCHI S.;
 RL SUBMITTED (JAN-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: INVOLVED IN THE REGULATION OF SECONDARY METABOLISM BY
 PHOSPHORYLATING, ON BOTH SER AND THR, THE AFSR GLOBAL REGULATORY
 CC PROTEIN INVOLVED IN THE CONTROL OF SECONDARY METABOLISM.
 CC -1- PTM: AUTOPHOSPHORYLATED AT THR AND SER RESIDUES.
 CC -1- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-
 PROTEIN KINASES.
 DR EMBL: D15062; E120137; -.
 DR EMBL: D45382; G642157; -.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 KW TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; ATP-BINDING;
 KW PHOSPHORYLATION.
 FT DOMAIN 16 271 PROTEIN KINASE.
 FT NP_BIND 22 30 ATP (BY SIMILARITY).
 FT BINDING 44 44 ATP (BY SIMILARITY).
 FT ACT_SITE 138 138 BY SIMILARITY.
 FT SEQUENCE 799 AA; 83781 MW; 545EE296 CRC32;
 SQ
 Query Match 32.3%; Score 65; DB 1; Length 799;
 Best Local Similarity 37.5%; Pred. No. 5.25e+00;
 Matches 9; Conservative 7; Mismatches 8; Indels 0; Gaps 0;
 Db 255 RNPADLQALPHLPFGSGDDSG 278
 QY 1 RCPGPGLOGLRQLQANGHAAG 24
 RESULT 13
 ID CYAA_SCHPO STANDARD; PRT; 1692 AA.
 AC P14605;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
 DE ADENYLATE CYCLASE (EC 4.6.1.1) (ATP PYROPHOSPHATE-LYASE) (ADENYLYL
 CYCLASE).
 DE CYCL.
 GN CYRI.
 OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
 CC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
 RN [1]
 RC SEQUENCE FROM N.A.
 RX MEDLINE: 90046723.
 RA YOUNG D., RIGGS M., FIELD J., VOJTEK A., BROEK D., WIGLER M.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 86:7989-7993(1989).
 RN [2]
 RC SEQUENCE FROM N.A.
 RX MEDLINE: 89345533.
 RA YAMAWAKI-KATAOKA Y., TAMAOKI T., CHOE H.-R., TANAKA H., KATAOKA T.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 86:5693-5697(1989).
 CC -1- FUNCTION: PLAYS ESSENTIAL ROLES IN REGULATION OF CELLULAR
 METABOLISM BY CATALYSING THE SYNTHESIS OF A SECOND MESSENGER,
 CAMP.
 CC -1- CATALYTIC ACTIVITY: ATP = 3', 5'-CYCLIC AMP + PYROPHOSPHATE.
 CC -1- IN CONTRAST TO YEAST CYCLASE, S.POMBE CYCLASE IS NOT LIKELY TO BE
 REGULATED BY RAS PROTEINS.
 CC -1- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN
 MANY PROTEINS. NUMBER IN THIS PROTEIN: 22.
 CC -1- SIMILARITY: BELONGS TO CLASS-3 OF ADENYLYL CYCLASES.

CC -1- SIMILARITY: CONTAINS A PP2C-LIKE DOMAIN.
 DR EMBL: M26699; G173339; -.
 DR EMBL: M24942; G173379; -.
 DR PIR: A33988; A33988.
 DR PIR: A33539; A33539.
 KW LYASE; REPEAT; LEUCINE-REPEAT; CAMP SYNTHESIS; MAGNESIUM.
 FT DOMAIN 326 999 LEUCINE-RICH REPEATS.
 FT REPEAT 326 349 LRR 1.
 FT REPEAT 350 396 LRR 2.
 FT REPEAT 397 420 LRR 3.
 FT REPEAT 421 444 LRR 4.
 FT REPEAT 445 466 LRR 5.
 FT REPEAT 467 492 LRR 6.
 FT REPEAT 493 517 LRR 7.
 FT REPEAT 518 540 LRR 8.
 FT REPEAT 541 563 LRR 9.
 FT REPEAT 564 587 LRR 10.
 FT REPEAT 588 616 LRR 11.
 FT REPEAT 617 674 LRR 12.
 FT REPEAT 675 698 LRR 13.
 FT REPEAT 699 722 LRR 14.
 FT REPEAT 723 744 LRR 15.
 FT REPEAT 745 798 LRR 16.
 FT REPEAT 799 822 LRR 17.
 FT REPEAT 823 852 LRR 18.
 FT REPEAT 853 892 LRR 19.
 FT REPEAT 893 922 LRR 20.
 FT REPEAT 923 950 LRR 21.
 FT REPEAT 951 999 LRR 22.
 FT DOMAIN 1008 1276 PP2C-LIKE.
 FT DOMAIN 1277 1692 CATALYTIC.
 SQ SEQUENCE 1692 AA; 190333 MW; 20E7D7D8 CRC32;
 Query Match 32.3%; Score 65; DB 1; Length 1692;
 Best Local Similarity 57.1%; Pred. No. 5.25e+00;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Db 1539 PVOLQGRRLERLKS 1552
 QY 4 PPGLOGLRQLQQA 17
 RESULT 14
 ID IL11_MOUSE STANDARD; PRT; 199 AA.
 AC P47873;
 DT 01-FEB-1996 (REL. 33, CREATED)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DE INTERLEUKIN-11 PRECURSOR (IL-11).
 GN IL11.
 OS MUS MUSCULUS (MOUSE).
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; RODENTIA.
 RN [1]
 RC SEQUENCE FROM N.A.
 RX MEDLINE: 97070356.
 RA MORRIS J.C., FINNERTY H., BENNET F., TURNER K.J., WOOD C.R.;
 RL EXP. HEMATOL. 24:1369-1376(1996).
 CC -1- FUNCTION: THIS PROTEIN STIMULATES PLASMACYTOMA PROLIFERATION,
 T-CELL-DEPENDENT DEVELOPMENT OF IMMUNOGLOBULIN-PRODUCING B
 CELLS AND SYNERGIZES WITH IL-3 IN SUPPORTING MURINE
 CC MEGAKARYOCYTE COLONY FORMATION (BY SIMILARITY).
 CC EMBL: U03421; G415654; -.
 DR MGD; MGI:107613; IL11.
 KW CYTOKINE; GROWTH FACTOR; SIGNAL.
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 199 INTERLEUKIN-11.
 SQ SEQUENCE 199 AA; 21522 MW; 1CB30772 CRC32;
 Query Match 31.8%; Score 64; DB 1; Length 199;
 Best Local Similarity 53.3%; Pred. No. 7.52e+00;
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

WIREH (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm

on: Thu Jul 30 09:35:23 1998; Maspar time 5.13 Seconds
230.058 Million cell updates/sec

ular output not generated.

Title: >US-08-938-548A-12
Description: (1-28) from US08938548A.pep
Perfect Score: 201
Sequence: 1 RGPGLGRLQLRLLQNGHAGILTM 28

Scoring table: PAM 150
Gap 15

Searched: 140542 seqs, 42109429 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrembls

1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phase 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

Statistics: Mean 30.024; Variance 52.279; scale 0.574

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	76	37.8	1174	4	Q95168 TIGHT JUNCTION PROTEIN	2.55e-01
2	75	37.3	1277	10	Q35821 PAR INTERACTING PROTEIN	3.68e-01
3	73	36.3	249	10	P97382 K+ CHANNEL BETA4 SUBUN	7.59e-01
4	69	34.3	331	12	Q91540 LEUCINE ZIPPER WITH BA	3.12e+00
5	68	33.8	2115	2	Q14980 NIMA PROTEIN.	4.42e+00
6	67	33.3	178	9	Q31955 YONC PROTEIN.	6.22e+00
7	67	33.3	413	10	Q35141 MAXP1.	6.22e+00
8	67	33.3	462	9	P95844 RNA POLYMERASE SIGMA F	6.22e+00
9	67	33.3	510	9	Q59913 RNA POLYMERASE SIGMA F	6.22e+00
10	67	33.3	514	9	P77951 RNA POLYMERASE SIGMA F	6.22e+00
11	67	33.3	525	9	Q59813 RNA POLYMERASE SIGMA F	6.22e+00
12	67	33.3	528	9	Q59814 RNA POLYMERASE SIGMA F	6.22e+00
13	67	33.3	568	3	Q27212 ARTICULIN P60.	6.22e+00
14	67	33.3	608	9	Q45998 DNA POLYMERASE III TAU	6.22e+00
15	66	32.8	130	9	Q60245 HRP.	8.74e+00
16	66	32.8	150	9	Q27642 DEOXYCYTIDINE-TRIPHOS	8.74e+00
17	66	32.8	253	9	Q54224 UNKNOWN GENE.	8.74e+00
18	66	32.8	304	3	Q22732 T24D5.1.	8.74e+00
19	66	32.8	319	9	Q05804 RNA-DIRECTED DNA POLYM	8.74e+00
20	66	32.8	335	12	Q91654 THYROID HORMONE INDUCE	8.74e+00

21	66	32.8	365	3	Q09636	HYPOTHETICAL 41.3 KD P	8.74e+00
22	66	32.8	377	9	Q50983	CARA.	8.74e+00
23	66	32.8	1465	3	Q17909	H06001.2.	8.74e+00
24	66	32.8	2591	9	Q54959	PRISTINAMYCIN I SYNTHA	8.74e+00
25	65	32.3	580	10	Q60811	RETINOID X RECEPTOR IN	1.22e+01
26	65	32.3	1344	10	Q35851	P160 MYB-BINDING PROTE	1.22e+01
27	65	32.3	2509	2	Q16702	FATTY ACID SYNTHASE (E	1.22e+01
28	64	31.8	140	9	Q50872	PUTATIVE SIGMA-54 DEPE	1.71e+01
29	64	31.8	263	6	Q31722	ORF263.	1.71e+01
30	64	31.8	288	9	Q51980	HOMODA HYDROLASE.	1.71e+01
31	64	31.8	294	9	Q53921	ORF328 (FRAGMENT).	1.71e+01
32	64	31.8	439	9	Q52495	HRPE.	1.71e+01
33	64	31.8	548	3	Q27932	PROTEIN-TYROSINE PHOSP	1.71e+01
34	64	31.8	549	2	Q99475	KM-102-DERIVED REDUCTA	1.71e+01
35	64	31.8	1348	3	Q04164	POTATIVE EPIDERMAL CEL	1.71e+01
36	63	31.3	556	9	P76237	FROM BASES 1860594 TO	2.37e+01
37	63	31.3	633	9	Q07309	NODQ.	2.37e+01
38	63	31.3	770	9	P73339	HYPOTHETICAL 85.5 KD P	2.37e+01
39	63	31.3	812	3	Q06452	EXTRACELLULAR MATRIX P	2.37e+01
40	63	31.3	881	2	Q15027	KIAA0310.	2.37e+01
41	63	31.3	972	11	P89521	POLYPROTEIN.	2.37e+01
42	63	31.3	1841	9	Q33958	TYLACTONE SYNTHASE MOD	2.37e+01
43	63	31.3	3654	9	Q30766	POLYKETIDE SYNTHASE MO	2.37e+01
44	62	30.8	527	12	Q91952	C-SRC TYROSINE KINASE.	3.29e+01
45	62	30.8	1690	2	Q12823	A TYPE IV COLLAGEN (FR	3.29e+01

ALIGNMENTS

RESULT 1
ID Q95168 PRELIMINARY; PRT; 1174 AA.
AC Q95168; 01-FEB-1997 (TREMBREL. 02, CREATED)
DT 01-FEB-1997 (TREMBREL. 02, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBREL. 05, LAST ANNOTATION UPDATE)
DE TIGHT JUNCTION PROTEIN.
GN ZO-2.
OS CANIS FAMILIARIS (DOG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; CARNIVORA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94179414.
RA JESAITIS L.A., GOODENOUGH D.A.;
RL J. CELL BIOL. 124:949-961(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96421547.
RA BEATCH M., JESAITIS L.A., GALLIN W., GOODENOUGH D.A.,
RA STEVENSON B.R.;
RL J. BIOL. CHEM. 271:25723-25726(1996).
RN [3]
RP SEQUENCE FROM N.A.
RA GOODENOUGH D.A.;
RL SUBMITTED (MAR-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [4]
RP SEQUENCE FROM N.A.
RA BEATCH M.;
RL SUBMITTED (AUG-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR ENBL; L27152; G1536970;
SQ SEQUENCE 1174 AA; 132085 MW; 2FA16B83 CRC32;

Query Match 37.8%; Score 76; DB 4; Length 1174;
Best Local Similarity 32.1%; Pred. No. 2.55e-01;
Matches 9; Conservative 12; Mismatches 7; Indels 0; Gaps 0;

Db 268 RSPSPRLGRPDHAGQSDSRPGVILM 295
QY 1 RGPGLGRLQLRLLQNGHAGILTM 28
RESULT 2
ID Q35821 PRELIMINARY; PRT; 1277 AA.

RA MASUDA S., MAUEL C., MEDIGUE C., MEDINA N., MELLADO R.P., MIZUNO M.,
RA MOESTL D., NAKAI S., NOBACK M., NOONE D., O'REILLY M., OGAWA K.,
RA OGIMURA A., ODEGA S., PARK S.H., PARRO V., POHL T.M., PORTETELLE D.,
RA PORWOLLIK S., PRESCOTT A.M., PRESECAN E., PUIC P., PURNELLE B.,
RA RAPOPORT G., REY M., REYNOLDS S., RIEGER M., RIVOLTA C., ROCHA E.,
RA ROCHE B., ROSE M., SADATE Y., SATO T., SCANLAN E., SCHLEICH S.,
RA SCHROETER R., SCOFFONE F., SEKIGUCHI J., SEKOWSKA A., SEROR S.J.,
RA SEROR P., SHIN B.S., SORDO B., SOROKIN A., TACCONI E., TAKAGI T.,
RA TAKAHASHI H., TAKEMARU K., TAKEUCHI M., TAMAKOSHI A., TANAKA T.,
RA TERPSTRA P., TOGNONI A., TOSATO V., UCHIYAMA S., WANDENBOL M.,
RA VANNIER F., VASSAROTTI A., VIARI A., WAMBUTT R., WEDLER H.,
RA WEITZENEGER T., WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K.,
RA YASUMOTO K., YATA K., YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E.,
RL NATURE 390:249-256(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.,
RA SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
EMBL: 299115; E1183561; -;
SEQUENCE 178 AA; 19609 MW; CA7B8ED0 CRC32;

Query Match 33.3%; Score 67; DB 9; Length 178;
Best Local Similarity 45.0%; Pred. No. 6.22e+00;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

Db 1 MATRQKALTEVGNHTGNL 20
QY 7 LQGLRQLQANGNHAAGIL 26
|||:| |||:| |

RESULT 7
ID O35141 PRELIMINARY; PRT; 413 AA.
AC O35141;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE MAX1.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RA SLENEV V.I., DE CAMILLI P.V.;
RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF002251; G2459833; -;
SQ SEQUENCE 413 AA; 46679 MW; 69A87FCC CRC32;

Query Match 33.3%; Score 67; DB 10; Length 413;
Best Local Similarity 69.2%; Pred. No. 6.22e+00;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 76 RVPVGLQQLRR 88
QY 1 RGPVGLQQLRR 13
||| ||||| |||:|

RESULT 8
ID P95644 PRELIMINARY; PRT; 462 AA.
AC P95644;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE RNA POLYMERASE SIGMA FACTOR.
GN RP0D.
OS RENIBACTERIUM SALMONINARUM.
OC EUBACTERIA; FIRMICUTES; ACTINOMYCETES; RENIBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC33209;
RA BARRIL A., CONCHA M.I., FIGUEROA J., LEON G.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.

CC -!- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC THEN IS RELEASED.
DR EMBL: Y10835; E294130; -;
DR PROSITE; PS00715; SIGMA70_1; 1.
DR PROSITE; PS00716; SIGMA70_2; 1.
KW TRANSCRIPTION REGULATION; SIGMA FACTOR; DNA-DIRECTED RNA POLYMERASE;
KW DNA-BINDING.
SQ SEQUENCE 462 AA; 50504 MW; E54BBF5 CRC32;

Query Match 33.3%; Score 67; DB 9; Length 462;
Best Local Similarity 31.8%; Pred. No. 6.22e+00;
Matches 7; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

Db 390 LQQLHSVLDLTSERAGVVM 411
QY 7 LQGLRQLQANGNHAAGILTM 28
|||:| |||:| |||:|

RESULT 9
ID Q59913 PRELIMINARY; PRT; 510 AA.
AC Q59913;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE RNA POLYMERASE SIGMA FACTOR.
GN HRDB.
OS STREPTOMYCES GRISEUS.
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; STREPTOMYCETACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IMEU 3570;
RX MEDLINE: 95189101.
RA MARCOS A.T., DIEZ B., GUTIERREZ S., FERNANDEZ F.J., OGUIZA J.A.,
RA MARTIN J.F.;
RL GENE 153:41-48(1995).

CC -!- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC THEN IS RELEASED.
DR EMBL: X75952; G581664; -;
DR PROSITE; PS00715; SIGMA70_1; 1.
DR PROSITE; PS00716; SIGMA70_2; 1.
KW TRANSCRIPTION REGULATION; SIGMA FACTOR; DNA-DIRECTED RNA POLYMERASE;
KW DNA-BINDING.
SQ SEQUENCE 510 AA; 55795 MW; C3CB64EB CRC32;

Query Match 33.3%; Score 67; DB 9; Length 510;
Best Local Similarity 31.8%; Pred. No. 6.22e+00;
Matches 7; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

Db 438 LQQLHSVLDLTSERAGVVM 459
QY 7 LQGLRQLQANGNHAAGILTM 28
|||:| |||:| |||:|

RESULT 10
ID P77951 PRELIMINARY; PRT; 514 AA.
AC P77951;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE RNA POLYMERASE SIGMA FACTOR.
GN HRDB.
OS STREPTOMYCES GRISEUS.
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; STREPTOMYCETACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RA SHINKAWA H., HATADA Y., OKADA M., KINASHI H., NIMI O.;
RL J. BIOCHEM. 118:494-499(1995).
CC -!- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC THEN IS RELEASED.
DR EMBL: L08071; G1617256; -;

Search completed: Thu Jul 30 09:35:47 1998
Job time : 24 secs.

WQ5RELH (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:27:56 1998; MasPar time 8.78 Seconds
589.674 Million cell updates/sec

ular output not generated.

Title: >US-08-938-548A-10
Description: (1-123) from US08938548A.pep
Perfect score: 899
Sequence: 1 VFWNAVTLILLLLPPALLS.....GRGCTVTVTALAPRGSGV 123

Scoring table: PAM 150
Gap 11

Searched: 140542 seqs, 42109429 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: spiremb15
1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phage 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

Statistics: Mean 39.425; Variance 86.061; scale 0.458

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Alt No.	Score	Query Match	Length	ID	Description	Pred. No.
1	106	11.8	679	8	023352 HYPOTHETICAL 74.6 KD P	2.41e-02
2	103	11.5	567	10	008770 PLATELET GLYCOPROTEIN	5.93e-02
3	98	10.9	641	10	008463 FRIZZLED PROTEIN HOMOL	2.57e-01
4	94	10.5	574	10	035298 ACYLOXYACYL HYDROLASE	8.09e-01
5	93	10.3	331	12	091640 LEUCINE ZIPPER WITH BA	1.07e+00
6	93	10.3	1174	4	095168 TIGHT JUNCTION PROTEIN	1.07e+00
7	92	10.2	106	8	041051 PROLINE- AND LEUCINE-R	1.42e+00
8	91	10.1	602	10	063821 CYCLOOXYGENASE ISOFORM	1.87e+00
9	91	10.1	602	10	063884 CYCLOOXYGENASE 1.	1.87e+00
10	90	10.0	833	2	099523 SORTILIN PRECURSOR	2.47e+00
11	89	9.9	602	10	062731 PROSTAGLANDIN H SYNTHA	3.25e+00
12	88	9.8	226	9	006319 HYPOTHETICAL 23.1 KD P	4.27e+00
13	88	9.8	335	12	091654 THYROID HORMONE INDUCE	4.27e+00
14	87	9.7	252	9	005582 HYPOTHETICAL 25.9 KD P	5.60e+00
15	87	9.7	583	12	090491 DNA BINDING PROTEIN EI	5.60e+00
16	87	9.7	599	2	015122 PROSTAGLANDIN ENDOPERO	5.60e+00
17	87	9.7	1238	3	018780 SIMILARITY TO MOUSE SM	5.60e+00
18	86	9.6	410	9	072841 HYPOTHETICAL 48.1 KD P	7.32e+00
19	86	9.6	1099	10	097527 NB-2.	7.32e+00
20	85	9.5	563	4	002839 PORCINE MEMBRANE COFAC	9.56e+00

21	85	9.5	878	11	083101 RNA-DEPENDENT RNA POLY	9.56e+00
22	84	9.3	84	10	060471 ANION EXCHANGER ISOFOR	1.25e+01
23	84	9.3	102	8	043495 108 PROTEIN PRECURSOR	1.25e+01
24	84	9.3	277	9	033285 HYPOTHETICAL 30.9 KD P	1.25e+01
25	84	9.3	701	11	065568 HYPOTHETICAL 72.6 KD P	1.25e+01
26	84	9.3	729	10	060470 ANION EXCHANGER 2 A (F	1.25e+01
27	84	9.3	1117	2	000542 DIACYLGLYCEROL KINASE	1.25e+01
28	83	9.2	147	10	061639 3' ORF.	1.62e+01
29	83	9.2	248	11	012396 NEUROVIRULENCE FACTOR.	1.62e+01
30	83	9.2	329	3	027916 MYOMODULIN (FRAGMENT)	1.62e+01
31	83	9.2	370	3	007974 MYOMODULIN PRECURSOR	1.62e+01
32	83	9.2	440	10	035849 LECITHIN:CHOLESTEROL A	1.62e+01
33	83	9.2	498	3	017021 MALTASE-LIKE PROTEIN A	1.62e+01
34	83	9.2	551	9	072405 PCBR.	1.62e+01
35	83	9.2	1109	4	019179 GUANYLATE CYCLASE E.	1.62e+01
36	82	9.1	264	9	095156 HYPOTHETICAL 27.7 KD P	2.10e+01
37	82	9.1	350	3	025413 MYOMODULIN NEUROPEPTID	2.10e+01
38	82	9.1	556	9	076237 FROM BASES 1860594 TO	2.10e+01
39	82	9.1	3122	11	089459 VERY LARGE TEGUMENT PR	2.10e+01
40	82	9.1	3729	9	033956 TYLACTONE SYNTHASE MOD	2.10e+01
41	81	9.0	40	9	046729 FERRIC ENTEROBACTIN TR	2.72e+01
42	81	9.0	347	9	076559 FROM BASES 2573751 TO	2.72e+01
43	81	9.0	734	3	016791 F18A12.4 PROTEIN.	2.72e+01
44	81	9.0	1110	4	002809 GUANYLATE CYCLASE ISOF	2.72e+01
45	81	9.0	1520	8	039160 MYOSIN.	2.72e+01

ALIGNMENTS

RESULT	1					
ID	023352	PRELIMINARY;	PRT;	679	AA.	
AC	023352;					
DT	01-JAN-1998	(TREMBLREL. 05, CREATED)				
DT	01-JAN-1998	(TREMBLREL. 05, LAST SEQUENCE UPDATE)				
DT	01-JAN-1998	(TREMBLREL. 05, LAST ANNOTATION UPDATE)				
DE	HYPOTHETICAL 74.6 KD PROTEIN					
OS	ARABIDOPSIS THALIANA (MOUSE-EAR CRESS)					
OC	EUKARYOTA; PLANTA; EMBRYOPHYTA; ANGIOSPERMAE; DICOTYLEDONEAE;					
ON	CAPPARALES; CRUCIFERAE.					
RP	SEQUENCE FROM N.A.					
RA	BEVAN M., STIEKEMA W., MURPHY G., WAMBUIT R., POHL T., TERRY N.,					
RA	KREIS M., KAVANAGH T., ENTIAN K.D., RIEGER M., JAMES R.,					
RA	FUIGMONNECH P., HATZOPOULOS P., OBERMAIER B., DUESTERHOFT A., JONES J.,					
RA	PAUME K., ANSORGE W., DELSENY M., BANCROFT I., MEWES H.W., SCHUELLER C.,					
RA	CHALWATZIS N.;					
RL	SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RA	EU ARABIDOPSIS SEQUENCING PROJECT, ESSA;					
RL	SUBMITTED (JUN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.					
DR	EMBL; 297337; E326841; -					
KW	HYPOTHETICAL PROTEIN.					
SQ	SEQUENCE 679 AA; 74635 MW; B301B713 CRC32;					

Query Match 11.8%; Score 106; DB 8; Length 679;
Best Local Similarity 32.3%; Pred. No. 2.41e-02;
Matches 20; Conservative 16; Mismatches 22; Indels 4; Gaps 4;

Db 598 KYCRSK-YEITHQNHNDNAADVLELAKREMPAELL-R-ASLRHTNEDQRFNLLNVRSA 654

QY 35 KTCSCRYELLHGAG-NHAAGILTLGRRRPPGPGLOGLQRLQRLQANGHAAAGILTMGRRA 93

Db 655 SP 656

QY 94 GA 95

RESULT 2

ID 008770 PRELIMINARY; PRT; 567

AC 008770;

DT 01-JUL-1997 (TREMBLREL. 04, CREATED)

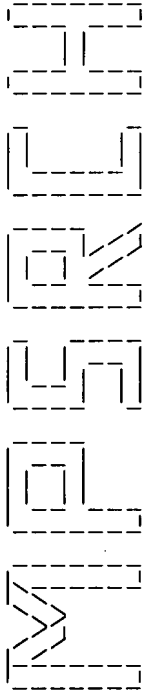
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)

Matches 19; Conservative 15; Mismatches 22; Indels 3; Gaps 3;
Db 119 PRLGQGNMAAV-SAEGN-GNRFVWIAVETSQL-PAPSPKDCSKVTFSQTQLRGIEV 174
QY 65 PPLGQRLQRLQANGNHAAGILTMGRRAGAELEPHPCSGRCCTVTTTALAPRGSGV 123

RESULT 15
ID Q90491 PRELIMINARY; PRT; 583 AA.
AC Q90491;
DT 01-NOV-1996 (TREMELREL. 01, CREATED)
DT 01-NOV-1996 (TREMELREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMELREL. 01, LAST ANNOTATION UPDATE)
DE DNA BINDING PROTEIN E12.
OS BRACHYDANTIO RERIO (ZEBRAFISH) (ZEBRA DANIO).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;
OC OSTEICHTHYES; ACTINOPTERYGII; CYPRINIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WT: TISSUE-WHOLE EMBRYO;
RY MEDLINE; 95001558.
WUELBECK C., FROMENTAL-RAMAIN C., CAMPOS-ORTEGA J.A.;
MECH. DEV. 46:73-85(1994).
DR EMBL; X76997; G509754; -.
KW DNA-BINDING.
SQ SEQUENCE 583 AA; 62625 MW; A874D9E2 CRC32;

Query Match 9.7%; Score 87; DB 12; Length 583;
Best Local Similarity 38.6%; Pred. No. 5.60e+00;
Matches 17; Conservative 10; Mismatches 15; Indels 2; Gaps 2;
Db 351 LLSAITSCHSAGAAALGSLRQAF-GLANRLPGLMSNHSEDAACL 393
QY 44 LLHGA-GNHAGILTLGKRPPGPGGLQRLQLLQANGNHAAGI 86

Search completed: Thu Jul 30 09:28:17 1998
Job time : 21 secs.



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:27:25 1998; MasPar time 5.12 Seconds
Popular output not generated. 602.279 Million cell updates/sec

Title: >US-08-938-548A-10
Description: (1-123) from US08938548A.pep
Perfect Score: 899
Sequence: 1 VPWAAVILLILLILLPPALLS.....GRGCPVTVTALAPRGSGV 123

Scoring table: PAM 150
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 41.408; Variance 75.349; scale 0.550

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	101	11.2	131	1	SECR_PIG SECRETIN PRECURSOR (FR	1.48e-02
2	96	10.7	566	1	FBLA_HUMAN FIBULIN-1, ISOFORM A P	7.91e-02
3	96	10.7	601	1	FBLB_HUMAN FIBULIN-1, ISOFORM B P	7.91e-02
4	96	10.7	683	1	FBLC_HUMAN FIBULIN-1, ISOFORM C P	7.91e-02
5	96	10.7	703	1	FBLD_HUMAN FIBULIN-1, ISOFORM D P	7.91e-02
6	91	10.1	440	1	LCAT_HUMAN PHOSPHATIDYLCHOLINE-ST	3.99e-01
7	89	9.9	1061	1	ANPA_HUMAN ATRIAL NATRIURETIC PEP	7.51e-01
8	88	9.8	254	1	41BL_HUMAN 4-1BB LIGAND (4-1BBL)	1.03e+00
9	88	9.8	438	1	LCAT_MOUSE PHOSPHATIDYLCHOLINE-ST	1.03e+00
10	86	9.6	322	1	YCEC_HAEN HYPOTHETICAL PROTEIN H	1.90e+00
11	86	9.6	440	1	LCAT_PAPAN PHOSPHATIDYLCHOLINE-ST	1.90e+00
12	86	9.6	497	1	SC14_YARLI SEC14 CYTOSOLIC FACTOR	1.90e+00
13	86	9.6	1663	1	CO3_RAT COMPLEMENT C3 PRECURSOR	1.90e+00
14	85	9.5	236	1	PLC1_BOVIN PLACENTAL LACTOGEN I P	2.57e+00
15	85	9.5	317	1	LIP1_PSVIM LIPASE 1 PRECURSOR (EC	2.57e+00
16	85	9.5	334	1	FEPD_ECOLI FERRIC ENTEROBACTIN TR	2.57e+00
17	85	9.5	1027	1	CAFE_RIFPA FIBRIL-FORMING COLLAG	2.57e+00
18	84	9.3	90	1	VGE_BPS13 E PROTEIN.	3.48e+00
19	84	9.3	90	1	VGE_BPPHX LYSIS PROTEIN.	3.48e+00
20	84	9.3	238	1	EFA3_HUMAN EPHRIN-A3 PRECURSOR (E	3.48e+00
21	84	9.3	251	1	C1OB_HUMAN COMPLEMENT C1Q SUBCOMP	3.48e+00
22	84	9.3	315	1	LIP3_MORSP LIPASE 3 PRECURSOR (EC	3.48e+00
23	84	9.3	319	1	YCEC_ECOLI HYPOTHETICAL 36.0 KD P	3.48e+00

24	84	9.3	384	1	RN_DROME GTPASE ACTIVATING PROT	3.48e+00
25	84	9.3	696	1	LSHR_PIG LUTROPIN-CHORIOGONADOT	3.48e+00
26	84	9.3	1001	1	PTPX_MOUSE PROTEIN-TYROSINE PHOSP	3.48e+00
27	84	9.3	2541	1	TALI_MOUSE TALIN.	3.48e+00
28	83	9.2	208	1	GPBB_PAPCY PLATELET GLYCOPROTEIN	4.69e+00
29	83	9.2	224	1	OXO2_HORVU OXALATE OXIDASE PRECUR	4.69e+00
30	83	9.2	235	1	FL3L_HUMAN SL CYTOKINE PRECURSOR	4.69e+00
31	83	9.2	248	1	ICP3_HSV11 INFECTED CELL PROTEIN	4.69e+00
32	83	9.2	252	1	ICP3_HSV1D INFECTED CELL PROTEIN	4.69e+00
33	83	9.2	263	1	ICP3_HSV1F INFECTED CELL PROTEIN	4.69e+00
34	83	9.2	440	1	LCAT_RABIT PHOSPHATIDYLCHOLINE-ST	4.69e+00
35	83	9.2	440	1	LCAT_RAT PHOSPHATIDYLCHOLINE-ST	4.69e+00
36	83	9.2	536	1	YABK_ECOLI HYPOTHETICAL 59.6 KD P	4.69e+00
37	83	9.2	615	1	ALBU_CHICK SERUM ALBUMIN PRECURSO	4.69e+00
38	83	9.2	628	1	FTSH_PORPU CELL DIVISION PROTEIN	4.69e+00
39	83	9.2	676	1	ICP0_HSVBJ TRANS-ACTING TRANSCRIP	4.69e+00
40	83	9.2	676	1	ICP0_HSVBK TRANS-ACTING TRANSCRIP	4.69e+00
41	83	9.2	1004	1	PTPX_RAT PROTEIN-TYROSINE PHOSP	4.69e+00
42	83	9.2	1310	1	ACE_RABIT ANGIOTENSIN-CONVERTING	6.30e+00
43	82	9.1	245	1	ICP3_HSVIN INFECTED CELL PROTEIN	6.30e+00
44	82	9.1	290	1	BLAI_XANMA METALLO-BETA-LACTAMASE	6.30e+00
45	82	9.1	1165	1	CYA6_MOUSE ADENYLATE CYCLASE, TYP	6.30e+00

ALIGNMENTS

RESULT ID	SECR_PIG	STANDARD;	PRT;	131 AA.
AC	P01279;			
DT	21-JUL-1986 (REL. 01, CREATED)			
DT	01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)			
DT	01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)			
DE	SECRETIN PRECURSOR (FRAGMENT).			
GN	SCT.			
OS	SUS SCROFA (PIG), BOS TAURUS (BOVINE), AND CAVIA PORCELLUS (GUINEA			
OC	PIG).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; ARTIODACTYLA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPECIES-PIG;			
RX	MEDLINE: 90192795.			
RA	KOPIN A.S., WHEELER M.B., LEITER A.B.;			
RL	PROC. NATL. ACAD. SCI. U.S.A. 87:2299-2303(1990).			
RN	[2]			
RP	SEQUENCE OF 30-56.			
RC	SPECIES-PIG;			
RX	MEDLINE: 70282334.			
RA	MUTT V., JORPES J.E., MAGNUSSON S.;			
RL	EUR. J. BIOCHEM. 15:513-519(1970).			
RN	[3]			
RP	SEQUENCE OF 30-59 AND 92-131.			
RC	SPECIES-PIG;			
RX	MEDLINE: 90370867.			
RA	CAFEVLIN G., JOERNVALL H., MUTT V.;			
RL	PROC. NATL. ACAD. SCI. U.S.A. 87:6781-6785(1990).			
RN	[4]			
RP	SYNTHESIS OF 30-131.			
RC	SPECIES-PIG;			
RA	BODANSZKY M., ONDETTI M.A., LEVINE S.D., NARAYANAN V.L., SALTZA M.V.,			
RL	SHEEHAN J.T., WILLIAMS N.J., SABO E.F.;			
RL	CHEM. IND. 1757-1758(1966).			
RN	[5]			
RP	SEQUENCE OF 30-56.			
RC	SPECIES-BOVINE;			
RX	MEDLINE: 81237102.			
RA	CARLOQUIST M., JOERNVALL H., MUTT V.;			
RL	FEBS LETT. 127:71-74(1981).			
RN	[6]			
RP	SEQUENCE OF 30-56.			
RC	SPECIES-C.PORCELLUS;			
RX	MEDLINE: 90254163.			
RA	BUSCAILL L., CAUVIN A., GOURLET P., GOSSEN D., DE NEEF P., RATHE J.,			

Query Match 10.78; Score 96; DB 1; Length 703;
Best Local Similarity 57.18; Pred.No. 7.91e-02;
Matches 16; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

Db 10 VPPLPGLLGGIALAAGVADVLLEACC 37
| | | | | | | | | | | | | | | | | |
QY 6 VTLLLLLLLPALLSLGVDQAQLPD-CC 32

RESULT 6
ID LCAT_HUMAN STANDARD; PRT; 440 AA.
AC P04180;
DT 20-MAR-1987 (REL. 04, CREATED)
DT 20-MAR-1987 (REL. 04, LAST SEQUENCE UPDATE)
DE 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE PHOSPHATIDYLCHOLINE-STEROL ACYLTRANSFERASE PRECURSOR (EC 2.3.1.43)
DE (LECITHIN-CHOLESTEROL ACYLTRANSFERASE) (PHOSPHOLIPID-CHOLESTEROL
ACYLTRANSFERASE).
LCAT
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 86205950.
RA MCLEAN J., FIELDING C., DRYANA D., DIEPLINGER H., BAER B., KOHR W.,
HENZEL W., LAWN R.;
RL PROC. NATL. ACAD. SCI. U.S.A. 83:2335-2339(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87091568.
RA MCLEAN J., WION K., DRYANA D., FIELDING C., LAWN R.;
RL NUCLEIC ACIDS RES. 14:9397-9406(1986).
RN [3]
RP SEQUENCE OF 17-440 FROM N.A.
RX MEDLINE; 88050946.
RA TATA F., CHAVES M.E., MARKHAM A.F., SCRACE G.D., WATERFIELD M.D.,
MCINTYRE N., WILLIAMSON R., HUMPHRIES S.E.;
RL BIOCHIM. BIOPHYS. ACTA 910:142-148(1987).
RN [4]
RP SEQUENCE OF 13-440 FROM N.A.
RX MEDLINE; 88049652.
RA ROGNE S., SKRETTING G., LARSEN F., MYKLEBOST O., MEVAG B.,
CARLSON L.A., HOLMQUIST L., GJONE E., PRYDZ H.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 148:161-169(1987).
RN [5]
RP PARTIAL SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE; 87137578.
RA YANG C., MANOOGIAN D., PAO Q., LEE F., KNAPP R.D., GOTTO A.M. JR.,
POWNALL H.J.;
RL J. BIOL. CHEM. 262:3086-3091(1987).
RN [6]
RP CARBOHYDRATE-BINDING SITES.
RX MEDLINE; 95338133.
RA SCHINDLER P.A., SETTNERI C.A., COLLET X., FIELDING C.J.,
BURLINGAME A.L.;
RL PROTEIN SCI. 4:791-803(1995).
RN [7]
RP VARIANT LEU-34.
RX MEDLINE; 92134273.
RA SKRETTING G., PRYDZ H.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 182:583-587(1992).
RN [8]
RP VARIANTS ILE-147 AND MET-371.
RX MEDLINE; 92147865.
RA KLEIN H.-G., LOHSE P., PRITCHARD P.H., BOJANOVSKI D., SCHMIDT H.,
BREWER H.B. JR.;
RL J. CLIN. INVEST. 89:499-506(1992).
RN [9]

FT CARBOHYD 535 535 POTENTIAL.
FT CARBOHYD 539 539 POTENTIAL.
SQ SEQUENCE 703 AA; 77274 MW; FOD06821 CRC32;

VARIANT ARG-170.
MEDLINE; 90316533.
RA TARAMELLI R., PONTIGLIO M., CANDIANI G., OTTOLENGHI S., DIEPLINGER H.,
CATAPANO A., ALBERS J., VERGANI C., MCLEAN J.;
RL HUM. GENET. 85:195-199(1990).
RN [10]
RP VARIANTS LYS-252 AND ILE-317.
RX MEDLINE; 92016763.
RA GOTODA T., YAMADA N., MURASE T., SAKUMA M., MURAYAMA N., SHIMANO H.,
KOZAKI K., ALBERS J., YAZAKI Y., AKANUMA Y.;
RL LANCET 338:778-781(1991).
RN [11]
RP VARIANT LYS-276.
RX MEDLINE; 92387377.
RA SKRETTING G., BLUMHOFF J.P., SOLHEIM J., PRYDZ H.;
RL FEBS LETT. 309:307-310(1992).
RN [12]
RP VARIANT ILE-317.
RX MEDLINE; 91315467.
RA MAEDA E., NAKA Y., MATOZAKI T., SAKUMA M., AKANUMA Y., YOSHINO G.,
KASUGA M.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 178:460-466(1991).
RN [13]
RP VARIANTS THR-117; TRP-159; CYS-182; PRO-233 AND MET-345.
RX MEDLINE; 93163362.
RA FUNKE H., VON ECKARDSTEIN A., PRITCHARD P.H., HORNBLY A.E.,
WIEBUSCH H., MOTTI C., HAYDEN M.R., DACHET C., JACOTOT B., GERDES U.,
FAERGEMAN O., ALBERS J.J., COLLEONI N., CATAPANO A., FROHLICH J.,
ASMANN G.;
RL J. CLIN. INVEST. 91:677-683(1993).
RN [14]
RP VARIANTS THR-117 AND CYS-182.
RX MEDLINE; 93305754.
RA HILL J.S., O.K., WANG X., PRITCHARD P.H.;
RL BIOCHIM. BIOPHYS. ACTA 1181:321-323(1993).
RN [15]
RP VARIANT HIS-164.
RX MEDLINE; 95331753.
RA STEYRER E., HAUBENWALLNER S., HOERL G., GIESSAU W., KOSTNER G.M.,
ZECHNER R.;
RL HUM. GENET. 96:105-109(1995).
RN [16]
RP VARIANTS ARG-57 AND 10 AA INSERTION IN POSITION 17.
RX MEDLINE; 95227171.
RA WIEBUSCH H., CULLEN P., OWEN J.S., COLLINS D., SHARP P.S., FUNKE H.,
ASMANN G.;
RL HUM. MOL. GENET. 4:143-145(1995).
RN [17]
RP VARIANT SER-54.
RX MEDLINE; 96400966.
RA OWEN J.S., WIEBUSCH H., CULLEN P., WATTS G.F., LIMA V.L.M.,
FUNKE H., ASMANN G.;
RL HUM. MUTAT. 8:79-82(1996).
CC -1- FUNCTION: CENTRAL ENZYME IN THE EXTRACELLULAR METABOLISM OF PLASMA
LIPOPROTEINS. AMONG OTHER SUBSTRATES IT ESTERIFIES THE FREE
CHOLESTEROL TRANSPORTED IN PLASMA LIPOPROTEINS.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + STEROL = STEROL ESTER +
1-ACYLGlycerophosphocholine (PALMITOYL, OLEOYL, AND LINOLEOYL CAN
BE TRANSFERRED; A NUMBER OF STEROLS, INCLUDING CHOLESTEROL, CAN
ACT AS ACCEPTOR).
CC -1- ENZYME REGULATION: APOLIPOPROTEIN A-I IS A POTENT ACTIVATOR FOR
THIS ENZYME.
CC -1- DISEASE: DEFECTS IN LCAT ARE THE CAUSE OF NORUM AND FISH EYE
DISEASES.
CC -1- SIMILARITY: PARTIAL WITH LIPASES (PANCREATIC, GASTRIC, HEPATIC,
LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
CC EMBL; M12625; G307117;
DR EMBL; X06537; -: NOT ANNOTATED_CDS.
DR EMBL; M26268; G187025;
DR EMBL; X04981; G34287; -;
DR EMBL; M17959; G386858; -;
DR PIR; A00571; XXHUN.
DR PIR; A29661; A29661.

CC LIPOPROTEINS. AMONG OTHER SUBSTRATES IT E
CC CHOLESTEROL TRANSPORTED IN PLASMA LIPOPRO

CC LIPOPROTEINS. AMONG OTHER SUBSTRATES IT E
CC CHOLESTEROL TRANSPORTED IN PLASMA LIPOPRO

FT DISULFID 1518 1590 BY SIMILARITY.
 FT DISULFID 1537 1661 BY SIMILARITY.
 FT THIOLEST 1010 1013
 FT CARBOHYD 939 939 PROBABLE.
 FT CARBOHYD 1617 1617 PROBABLE.
 FT CONFLICT 721 722 LK -> KL (IN REF. 2).
 SQ SEQUENCE 1663 AA; 186460 MW; 0428CF63 CRC32;

Query Match 9.6%; Score 86; DB 1; Length 1663;
 Best Local Similarity 66.7%; Pred. No. 1.90e+00;
 Matches 10; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 10 LVLLLLASSLLALG 24
 I:|||||:::|||||
 QY 8 LVLLLLPALLSLG 22

RESULT 14
 ID PLCL_BOVIN STANDARD; PRT; 236 AA.
 AC P09611;
 01-MAR-1989 (REL. 10, CREATED)
 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE PLACENTAL LACTOGEN I PRECURSOR (BPLP-I).
 GN PLI.
 OS BOS TAURUS (BOVINE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; ARTIODACTYLA.
 RN RN
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 90256825.
 RA YAMAKAWA M., TANAKA M., KOYAMA M., KAGESATO Y., WATAHAKI M.,
 RA YAMAMOTO M., NAKASHIMA K.;
 RL J. BIOL. CHEM. 265:8915-8920(1990).
 RN [2]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE: 89207456.
 RA SCHULER L.A., SHIMOMURA K., KESSLER M.A., ZIELER C.G., BREMEL R.D.;
 RL BIOCHEMISTRY 27:8443-8448(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 91166732.
 RA KESSLER M.A., SCHULER L.A.;
 RL DNA CELL BIOL. 10:93-101(1991).
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE SOMATOTROPIN/PROLACTIN FAMILY.
 DR EMBL: J02840; G163336; -
 DR EMBL: M33268; G163629; -
 DR EMBL: M65218; G163279; ALT_SEQ.
 DR EMBL: M65219; G163281; ALT_SEQ.
 DR EMBL: M65224; G163277; -
 DR EMBL: M65220; G163277; JOINED.
 DR EMBL: M65221; G163277; JOINED.
 DR EMBL: M65222; G163277; JOINED.
 DR EMBL: M65223; G163277; JOINED.
 DR PIR: A36284; A36284.
 DR PROSITE: PS00266; SOMATOTROPIN_1; 1.
 DR PROSITE: PS00338; SOMATOTROPIN_2; 1.
 KW HORMONE; PLACENTA; GLYCOPROTEIN; SIGNAL.
 FT SIGNAL 1 36
 FT CHAIN 37 236
 FT DISULFID 98 214 PLACENTAL LACTOGEN.
 FT DISULFID 231 236 BY SIMILARITY.
 FT CARBOHYD 89 89 BY SIMILARITY.
 FT CONFLICT 37 37 PROBABLE.
 FT CONFLICT 94 94 V -> A (IN REF. 2).
 FT CONFLICT 94 94 V -> F (IN REF. 2 AND 3).
 FT CONFLICT 170 170 V -> M (IN G163281).
 SQ SEQUENCE 236 AA; 26908 MW; 88029761 CRC32;

Query Match 9.5%; Score 85; DB 1; Length 236;
 Best Local Similarity 38.2%; Pred. No. 2.57e+00;
 Matches 13; Conservative 8; Mismatches 11; Indels 2; Gaps 2;

Db 22 LVLLLVSNLLCQGVEDYA-P-YCKNQPCNCRI 53
 I:|||||:::|||||
 QY 8 LVLLLLPALLSLGVDQAQPLDCCCRQKTCSCRL 41

RESULT 15
 ID LIPI_PSVIM STANDARD; PRT; 317 AA.
 AC Q02104;
 DT 01-JUL-1993 (REL. 26, CREATED)
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE LIPASE 1 PRECURSOR (BC 3.1.1.3) (TRIACYLGLYCEROL LIPASE).
 GN LIPI.
 OS PSYCHROBACTER IMMOBILIS.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
 CC NEISSERIACEAE.
 RN RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-B10;
 RX MEDLINE: 93144351.
 RA ARPAGNY J.L., FELLER G., GERDAY C.;
 RL BIOCHIM. BIOPHYS. ACTA 1171:331-333(1993).
 CC -1- FUNCTION: CATALYZES LIPOLYSIS AT TEMPERATURES AS LOW AS 3 DEGREES
 CELSIUS.
 CC -1- CATALYTIC ACTIVITY: TRIACYLGLYCEROL + H(2)O = DIACYLGLYCEROL +
 A FATTY ACID ANION.
 CC EMBL: X67712; G747876; -
 DR PIR: S26486; S26486.
 DR PIR: S28225; S28225.
 DR PROSITE: PS00120; LIPASE_SER; FALSE_NEG.
 KW HYDROLASE; LIPID DEGRADATION; SIGNAL.
 FT SIGNAL 1 27
 FT CHAIN 28 317
 FT ACT_SITE 74 74 LIPASE I.
 FT ACT_SITE 142 142 POTENTIAL.
 FT ACT_SITE 142 142 CHARGE RELAY SYSTEM (BY SIMILARITY).
 SQ SEQUENCE 317 AA; 35251 MW; 50684A86 CRC32;

Query Match 9.5%; Score 85; DB 1; Length 317;
 Best Local Similarity 35.1%; Pred. No. 2.57e+00;
 Matches 13; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

Db 95 LIIPDLGFGESKPMASDYRSEARLRLHLLQAKG 131
 I:|||||:::|||||
 QY 13 LIIPALLSLGVDQAQPLDCCCRQKTCSCRIYELHAG 49

Search completed: Thu Jul 30 09:27:36 1998
 Job time : 11 secs.

WIREH

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:26:50 1998; Maspar time 7.38 Seconds
608.699 Million cell updates/sec
Linear output not generated.

Title: >US-08-938-548A-10
Description: (1-123) from US08938548A.pep
Perfect Score: 899
Sequence: 1 VPWAVVTLTLLLLPPALLS.....GRCPTVTTLAPRGSGV 123

Scoring table: PAM 150
Gap 11
Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir56
1:pir1 2:pir2 3:pir3 4:pir4 5:nrl3d

Statistics: Mean 39.443; Variance 86.453; scale 0.456

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	101	11.2	131	1	SEPG secretin precursor -	9.71e-02
2	98	10.9	641	2	probable intercellula	2.32e-01
3	96	10.7	601	2	fibulin 1 precursor,	4.12e-01
4	96	10.7	683	2	fibulin 1 precursor,	4.12e-01
5	92	10.2	838	2	eyeless, long form -	1.27e+00
6	91	10.1	440	1	XXHUN phosphatidylcholine--	1.68e+00
7	91	10.1	602	2	cyclooxygenase 1 - ra	1.68e+00
8	89	9.9	602	2	prostaglandin G/H syn	2.91e+00
9	89	9.9	1061	1	oxyHAR natriuretic peptide r	2.91e+00
10	88	9.8	245	2	FLT3/FLK2 ligand (clo	3.81e+00
11	88	9.8	438	1	XXMSN phosphatidylcholine--	3.81e+00
12	88	9.8	491	2	stromelysin 3 (EC 3.4	3.81e+00
13	87	9.7	593	2	DNA binding protein E	4.99e+00
14	87	9.7	599	2	A36746 prostaglandin-endoper	4.99e+00
15	86	9.6	322	2	G64151 hypothetical protein	6.52e+00
16	86	9.6	410	2	J47705 hypothetical protein	6.52e+00
17	86	9.6	440	2	JC1502 phosphatidylcholine--	6.52e+00
18	86	9.6	497	2	S43745 phosphatidylinositol-	6.52e+00
19	86	9.6	1663	1	C3RT complement C3 precurs	6.52e+00
20	85	9.5	236	2	A37930 placental lactogen pr	8.50e+00
21	85	9.5	317	2	S28225 triacylglycerol lipas	8.50e+00
22	85	9.5	317	2	S57275 triacylglycerol lipas	8.50e+00
23	85	9.5	334	2	S16296 ferric enterobactin t	8.50e+00

24	85	9.5	1027	2	S28774 collagen alpha chain	8.50e+00
25	84	9.3	90	1	ZEBPF4 gene E protein - phag	1.11e+01
26	84	9.3	91	2	S47060 lysis protein - phage	1.11e+01
27	84	9.3	91	2	JS0455 gene E protein - phag	1.11e+01
28	84	9.3	102	2	S26409 protein 108 precursor	1.11e+01
29	84	9.3	253	1	CIHQUB complement subcomponent	1.11e+01
30	84	9.3	315	2	S14276 triacylglycerol lipas	1.11e+01
31	84	9.3	319	2	C64852 hypothetical 36.0K pr	1.11e+01
32	84	9.3	329	2	B41344 lutropin-choriogonad	1.11e+01
33	84	9.3	329	2	D41344 lutropin-choriogonad	1.11e+01
34	84	9.3	331	2	C41344 lutropin-choriogonad	1.11e+01
35	84	9.3	333	2	B48122 GTPase-activating pro	1.11e+01
36	84	9.3	384	2	A48122 lutropin-choriogonad	1.11e+01
37	84	9.3	696	2	A41344 lutropin-choriogonad	1.11e+01
38	84	9.3	701	2	S61239 hypothetical protein	1.11e+01
39	84	9.3	2541	2	S11661 talin - mouse	1.11e+01
40	83	9.2	498	2	S55362 maltase-like protein	1.44e+01
41	83	9.2	615	1	ABCHS serum albumin precurs	1.44e+01
42	83	9.2	628	2	S73248 hypothetical protein	1.44e+01
43	83	9.2	676	1	EDBE23 immediate-early prote	1.44e+01
44	83	9.2	676	1	EDBE22 immediate-early prote	1.44e+01
45	83	9.2	1309	2	S35484 peptidyl-dipeptidase	1.44e+01

ALIGNMENTS

RESULT ENTRY	1	SEPG	#type complete
TITLE	secretin precursor - pig		
ORGANISM	#formal_name Sus scrofa domestica	#common_name domestic pig	
DATE	24-Apr-1984	#sequence_revision 12-Apr-1996	#text_change
ACCESSIONS	20-Mar-1998		
REFERENCE	B35094; A01544; A36052		
#authors	A35094		
#journal	Kopin, A.S.; Wheeler, M.B.; Leiter, A.B.		
#title	Proc. Natl. Acad. Sci. U.S.A. (1990) 87:2299-2303		
#cross-references	Secretin: structure of the precursor and tissue distribution of the mRNA.		
#accession	B35094		
#molecule_type	mRNA		
#residues	1-131	#label KOP	
#cross-references	GB:M31496; NID:g164670; PID:g164671		
REFERENCE	A31147		
#authors	Mutt, V.; Jorpes, J.E.; Magnusson, S.		
#journal	Eur. J. Biochem. (1970) 15:513-519		
#title	Structure of porcine secretin. The amino acid sequence.		
#cross-references	MUID:70282334		
#accession	A01544		
#molecule_type	protein		
#residues	30-56	#label MUT	
#note	tryptic peptides were sequenced		
REFERENCE	A36052		
#authors	Gafvelin, G.; Joernvall, H.; Mutt, V.		
#journal	Proc. Natl. Acad. Sci. U.S.A. (1990) 87:6781-6785		
#title	Processing of prosecretin: isolation of a secretin precursor from porcine intestine.		
#cross-references	MUID:90370867		
#accession	A36052		
#status	preliminary		
#molecule_type	protein		
#residues	30-59, R', 92-131	#label GAF	
REFERENCE	A30916		
#authors	Bodanszky, M.; Ondetti, M.A.; Levine, S.D.; Narayanan, V.L.; Saltza, M.V.; Sheehan, J.T.; Williams, N.J.; Sabo, E.F.		
#journal	Chem. Ind. (1966) 1757-1758		
#title	Synthesis of a heptacosapeptide amide with the hormonal activity of secretin.		
#contents	annotation		
#note	synthesis confirmed the proposed structure of the natural hormone		
CLASSIFICATION	#superfamily glucagon		
KEYWORDS	amidated carboxyl end; duodenal mucosa; duplication; hormone;		

```

#molecule_type mRNA
#residues 13-440 #label ROG
##cross-references GB:M17959; NID:g187026; PID:g386858
REFERENCE
#authors Tata, F.; Chaves, M.E.; Markham, A.F.; Scrase, G.D.; Waterfield, M.D.; McIntyre, N.; Williamson, R.; Humphries, S.E.
#journal Biochim. Biophys. Acta (1987) 910:142-148
#title The isolation and characterisation of cDNA and genomic clones for human lecithin:cholesterol acyltransferase.
#cross-references MUID:88050946
#accession J00036
#molecule_type mRNA
#residues 17-256,'H',258-440 #label TAT
##cross-references GB:X06537; NID:g34284; GB:M26269; NID:g187024; PID:g187025
##note the authors translated the codon CAT for residue 241 as Ile and CAG for residues 251, 304, 368, 373, and 384 as Leu
A29133
REFERENCE
#authors Yang, C.; Mancoogian, D.; Pao, Q.; Lee, F.; Knapp, R.D.; Gatto, Jr., A.M.; Pownall, H.J.
#journal J. Biol. Chem. (1987) 262:3086-3091
#title Lecithin:cholesterol acyltransferase. Functional regions and a structural model of the enzyme.
#cross-references MUID:87137578
#accession A29133
#molecule_type protein
#residues 25-284,'Q',286-333,'Q',335-440 #label YAN
REFERENCE
#authors Bujo, H.; Kusunoki, J.; Ogasawara, M.; Yamamoto, T.; Ohta, Y.; Shimada, T.; Saito, Y.; Yoshida, S.
#journal Biochem. Biophys. Res. Commun. (1991) 181:933-940
#title Molecular defect in familial lecithin:cholesterol acyltransferase (LCAT) deficiency: a single nucleotide insertion in LCAT gene causes a complete deficient type of the disease.
#cross-references MUID:92109783
#accession I52260
#status translated from GB/EMBL/DBJ
#molecule_type DNA
#residues 25-34,'AHHAQG' #label BUJ
##cross-references GB:S74079; NID:g241428; PID:g241429
##note defective frame shift mutant sequence
A57914
REFERENCE
#authors Schandler, P.A.; Settineri, C.A.; Collet, X.; Fielding, C.J.; Burlingame, A.L.
#journal Protein Sci. (1995) 4:791-803
#title Site-specific detection and structural characterization of the glycosylation of human plasma proteins
#description lecithin:cholesterol acyltransferase and apolipoprotein D using HPLC/electrospray mass spectrometry and sequential glycosidase digestion.
#contents annotation; peptide sequences; N- and O-glycosylation
COMMENT Apolipoprotein A-I (see PIR:LPHUA1) is a potent activator of this enzyme.
GENECS
#gene GDB:LCAT
##cross-references GDB:119359; OMIM:245900
#map_position 16q22.1-16q22.1
FUNCTION
#description catalyzes the transfer of sn-2 fatty acyl groups from phosphatidylcholine (lecithin) to sterol to form sterol fatty esters and 1-acylglycerophosphocholine
#note palmitoyl, oleoyl, and linoleoyl residues can be transferred; a number of sterols, including cholesterol, can act as acceptor
#superfamily phosphatidylcholine--sterol acyltransferase
#acyltransferase; cholesterol; glycoprotein; lipid metabolism; lipoprotein
CLASSIFICATION
KEYWORDS
#domain signal sequence #status predicted #label SIG\
1-24
#product phosphatidylcholine--sterol acyltransferase
25-440
FEATURE

```



```

Db 14 LLLLLPPLLLLRGSHAGNL 35
QY 8 LLLLLPPALLSL-GVDAQPL 28

RESULT 10
ENTRY S43293 #type complete
TITLE FLT3/FLK2 ligand (clone S109) - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 20-Oct-1994 #sequence_revision 10-Nov-1995 #text_change
10-Nov-1995
ACCESSIONS S43293
REFERENCE S43290
#authors Hannun, C.; Culpepper, J.; Campbell, D.; McClanahan, T.;
Zurawski, S.; Bazan, J.F.; Kastelein, R.; Hudak, S.;
Wagner, J.; Mattson, J.; Luh, J.; Duda, G.; Martins, N.;
Peterson, D.; Menon, S.; Shanafelt, A.; Muench, M.; Kelner,
G.; Nankawa, R.; Rennick, D.; Roncarolo, M.G.; Zlotnik,
A.; Rosnet, O.; Dubreuil, P.; Birnbaum, D.; Lee, F.
#journal Nature (1994) 368:643-648
#title Ligand for FLT3/FLK2 receptor tyrosine kinase regulates
growth of haematopoietic stem cells and is encoded by
variant RNAs.
#accession S43293
#status preliminary
#molecule_type mRNA
#residues 1-245 #label HAN
#note the authors translated the codon AGT for residue 25 as
Met
SUMMARY #length 245 #molecular-weight 27404 #checksum 295
Query Match 9.88; Score 88; DB 2; Length 245;
Best Local Similarity 50.08; Pred. No. 3.81e+00;
Matches 10; Conservative 8; Mismatches 1; Indels 1; Gaps 1;
Db 7 AWSPTNLLLLLLSSGLMG 26
QY 2 PAAVT-LLLLLLPPALLS 20

RESULT 11
ENTRY XXSN #type complete
TITLE phosphatidylcholine--sterol O-acyltransferase (EC 2.3.1.43)
ALTERNATE_NAMES precursor - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
05-Sep-1997
ACCESSIONS A34158; S21370
REFERENCE A34158
#authors Warden, C.H.; Langner, C.A.; Gordon, J.I.; Taylor, B.A.;
McLean, J.W.; Lusa, A.J.
#journal J. Biol. Chem. (1989) 264:21573-21581
#title Tissue-specific expression, developmental regulation, and
chromosomal mapping of the lecithin:cholesterol
acyltransferase gene. Evidence for expression in brain and
testes as well as liver.
#cross-references MUID:90094326
#accession A34158
#molecule_type mRNA
#residues 1-438 #label WAR
#cross-references GB:J05154; NID:g198759; PID:g293697
#note the authors translated the codon ATG for residue 411 as
Leu
REFERENCE S21370
#authors Meroni, G.; Maggaretti, N.; Magnaghi, P.; Taramelli, R.
#submission submitted to the EMBL Data Library, July 1990
#description Promoter and 5' flanking sequences of the mouse LCAT gene.
#accession S21370
#molecule_type DNA
#residues 1-14 #label MER

##cross-references EMBL:X54095; NID:g52873; PID:g52874
COMMENT The active enzyme catalyzes the transfer of acyl groups from
lecithin to sterol to form sterol esters. Palmitoyl, oleoyl, and
linoleoyl residues can be transferred; a number of sterols,
including cholesterol, can act as acceptor. Apolipoprotein A-I is
a potent activator for this enzyme.
CLASSIFICATION #superfamily phosphatidylcholine--sterol acyltransferase
KEYWORDS acyltransferase; glycoprotein; lipid metabolism; lipoprotein
FEATURE
1-24 #domain signal sequence #status predicted #label SIG\
25-438 #product phosphatidylcholine--sterol acyltransferase
#status predicted #label MAT\
44,108,296,408 #binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY #length 438 #molecular-weight 49765 #checksum 1794
Query Match 9.88; Score 88; DB 1; Length 438;
Best Local Similarity 75.08; Pred. No. 3.81e+00;
Matches 12; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Db 7 PQWRVLLLLGLLPPA 22
QY 2 PAAVTLLLLLLPPA 17

RESULT 12
ENTRY JC6197 #type complete
TITLE stromelysin 3 (EC 3.4.24.-) - rat
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 11-Apr-1997 #sequence_revision 09-May-1997 #text_change
20-Mar-1998
ACCESSIONS JC6197
REFERENCE JC6197
#authors Okada, A.; Saez, S.; Misumi, Y.; Basset, P.
#journal Gene (1997) 185:187-193
#title Rat stromelysin 3: cDNA cloning from healing skin wound,
activation by furin and expression in rat tissues.
#contents Skin wounds
#accession JC6197
#molecule_type mRNA
#residues 1-491 #label OKA
#cross-references GB:U46034
COMMENT This protein is a member of the matrix metalloproteinase family.
CLASSIFICATION #superfamily interstitial collagenase; hemoexin repeat
homology; matrix metalloproteinase homology
KEYWORDS hydrolase; metalloproteinase; zinc; zymogen
FEATURE
52-261 #domain matrix metalloproteinase homology #label MMP\
291-483 #domain hemoexin repeat homology #label PXN\
84,218,222,228 #binding_site zinc, catalytic (Cys, His, His, His)
218,222,228 #binding_site zinc, catalytic (His) (active) #status
predicted\
219 #active_site Glu #status predicted
SUMMARY #length 491 #molecular-weight 55511 #checksum 284
Query Match 9.88; Score 88; DB 2; Length 491;
Best Local Similarity 35.08; Pred. No. 3.81e+00;
Matches 21; Conservative 16; Mismatches 18; Indels 5; Gaps 3;
Db 23 LLLLLLPPQLMA---RARP-PENHRHPVK-RYPQLPALPNLSVPASHWVPGPAS 77
QY 8 LLLLLLPPALLSLGVDAQPLPDCRCRKTCSCRLYELLHGAGNAAAGTLTGKRRPGPG 67

RESULT 13
ENTRY I50518 #type complete
TITLE DNA binding protein E12 - zebra fish
ORGANISM #formal_name Brachydanio rerio #common_name zebra fish
DATE 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
16-Feb-1997
ACCESSIONS I50518; S49141
REFERENCE I50518

```

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Run on: Thu Jul 30 09:26:03 1998; MasPar time 5.63 Seconds
Molecular output not generated. 335.050 Million cell updates/sec

Title: >US-08-938-548A-10
Description: (1-123) from US08938548A.pep
Perfect Score: 899
Sequence: 1 VPWAAVTLVLLLLLPALLS.....GRGCPVTVTALAPRGSGV 123

Scoring table: PAM 150
Gap 11

Searched: 124785 seqs, 15338987 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq31-2
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27

Statistics: Mean 29.422; Variance 127.873; scale 0.230

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	96	10.7	566	2 R11148	Fibulin A.	5.13e+00
2	96	10.7	601	2 R11149	Fibulin B.	5.13e+00
3	96	10.7	683	2 R11150	Fibulin C.	5.13e+00
4	91	10.1	440	2 P70134	Natural recombinant h	1.21e+01
5	91	10.1	440	25 W24789	Human lecithin-choles	1.21e+01
6	88	9.8	254	12 R64190	Human 4-IBB-L polypep	2.01e+01
7	88	9.8	254	25 W26557	Human 4-IBB ligand.	2.01e+01
8	87	9.7	111	7 R35515	Tryptophan aporepress	2.38e+01
9	87	9.7	599	4 R21690	Prostaglandin endoper	2.38e+01
10	85	9.5	363	23 W12414	Porcine complement in	3.33e+01
11	84	9.3	35	1 P94256	Truncated E protein f	3.93e+01
12	84	9.3	234	16 R82605	Eph transmembrane tyr	3.93e+01
13	84	9.3	238	13 R71481	Human hex-L protein.	3.93e+01
14	84	9.3	1239	9 R45945	Glutamic acid recepto	3.93e+01
15	83	9.2	179	8 R50056	ICP34.5 fragment.	4.63e+01
16	83	9.2	235	13 R66175	Human s86/S109 Flt3 1	4.63e+01
17	83	9.2	235	12 R67541	Human flt-3 ligand.	4.63e+01
18	83	9.2	222	4 R21708	HSV-1 (CVG-2) ICP34.5	4.63e+01
19	83	9.2	238	8 R50049	ICP34.5 fragment.	4.63e+01

20	83	9.2	263	4 R21706	HSV-1 (F) ICP34.5 pro	4.63e+01
21	83	9.2	338	4 R21707	HSV-1 (MGH-10) ICP34.	4.63e+01
22	83	9.2	551	13 R77858	S. clavuligerus ORF1	4.63e+01
23	82	9.1	35	1 P91354	Amino acid sequence o	5.46e+01
24	82	9.1	126	8 R50054	ICP34.5 fragment.	5.46e+01
25	82	9.1	3729	25 W22603	Tylosone synthase OR	5.46e+01
26	81	9.0	152	27 W29307	Wild-type avidin prot	6.44e+01
27	81	9.0	190	12 R60615	Human PDGF-B 109 subu	6.44e+01
28	81	9.0	226	2 R22673	v-sis protein p8sis.	6.44e+01
29	81	9.0	241	1 P80597	CV-sis gene encoded p	6.44e+01
30	81	9.0	241	8 R40967	CV-sis gene product.	6.44e+01
31	81	9.0	241	12 R63472	Recombinant platelet	6.44e+01
32	81	9.0	271	11 P80595	Recombinant platelet	6.44e+01
33	81	9.0	271	8 R40963	PDGF Bv-sis.	6.44e+01
34	81	9.0	271	12 R63468	Recombinant platelet	6.44e+01
35	81	9.0	282	12 R60616	Human PDGF-B 119 link	6.44e+01
36	80	8.9	226	3 R13757	Proactin.	7.58e+01
37	80	8.9	492	5 R24863	Sequence of pre-pro s	7.58e+01
38	80	8.9	521	10 R54988	Sweet potato feathery	7.58e+01
39	80	8.9	684	22 W15287	Human alpha3(IX) coll	7.58e+01
40	80	8.9	969	24 W25170	Human insulinoma-asso	7.58e+01
41	80	8.9	986	24 W25171	Human insulinoma-asso	7.58e+01
42	80	8.9	1012	26 W35296	Macaque islet cell an	7.58e+01
43	80	8.9	1015	27 W35345	Human protein tyrosin	7.58e+01
44	80	8.9	1015	23 W18092	Type I diabetes-assoc	7.58e+01
45	80	8.9	1337	14 R85203	hudep-1.	7.58e+01

ALIGNMENTS

RESULT 1	
ID R11148 standard; Protein; 566 AA.	
AC R11148;	
DE 21-MAY-1991 (first entry)	
DT Fibulin A.	
KW Beta-1 integrin; adhesion; receptor; fibronectin.	
OS Homo sapiens.	
EH Key	Location/Qualifiers
FT peptide	1..29
FT protein	30..566
FT modified_site	/label= fibulin A
FT modified_site	/label= N-linked glycosylation
FT modified_site	535
FT modified_site	/label= N-linked glycosylation
FT modified_site	339
FT region	/label= N-linked glycosylation
FT region	36..144
FT region	/label= type I motif
FT region	36..69
FT region	/label= repeat unit 1
FT region	112..144
FT region	/label= repeat unit 2
FT region	144..179
FT region	/label= Glu/Asp-rich region
FT region	180..566
FT region	/label= type II motif
FT region	180..214
FT region	/label= repeat unit 1
FT region	215..219
FT region	/label= consensus pentapeptide
FT region	220..260
FT region	/label= repeat unit 2
FT region	261..265
FT region	/label= consensus pentapeptide
FT region	266..306
FT region	/label= repeat unit 3
FT region	307..311
FT region	/label= consensus pentapeptide
FT region	312..354
FT region	/label= repeat unit 4
FT region	355..359

CC the repeat lacking two Cys residues. The same motif is found in
 CC complement components C3a, C4a and C5a; the inverse is found in
 CC albumin, vitamin D-binding protein and alpha-fetoprotein. The
 CC disulphide-stabilized loop structure is thought to be conserved.
 CC The type II motif, related to repeats found in epidermal growth
 CC factor precursor is a 6-Cys pattern repeated nine times, although
 CC the ninth repeat in the A form is incomplete. Four of the repeats,
 CC (2, 3, 4 and 9) differ from the EGF-like motif in having a 4-6
 CC residue insertion between cysteines 4 and 5 instead of the usual
 CC single residue. Embodied within repeats 5, 6, 7 and 8 is the
 CC consensus sequence for Asp and Asn hydroxylation. The 7th repeat
 CC contains the consensus O-glycosylation sequence, CXCPX.
 CC Immediately following each repeat is a pentapeptide with the
 CC consensus sequence XD(I/V)(D/N)E. Fibulin binds to the cyto-
 CC plasmic domain of the beta1 subunit of integrin adhesion
 CC receptors in a cation-dependent, EDTA-reversible manner. It can
 CC be used to manipulate adhesion of cells to fibronectin, collagen,
 CC laminin, and possibly also other proteins. Antibodies reactive
 CC with the protein have important diagnostic and therapeutic uses.
 CC See also R1147, R1148 and R1150.
 CC Sequence 601 AA;

Query Match 10.7%; Score 96; DB 2; Length 601;
 Best Local Similarity 57.1%; Pred. No. 5.13e+00;
 Matches 16; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

Db 10 vplllllglllaagvdadvleacc 37
 QY 6 VTLLLLLLPALLSLGVDQAQLPD-CC 32

RESULT 3
 ID R1150 standard; Protein; 683 AA.
 AC R1150;
 DT 21-MAY-1991 (first entry)
 DE Fibulin C.
 KW Beta-1 integrin; adhesion; receptor; fibronectin.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT peptide 1..29
 FT protein /label= signal sequence
 FT protein /label= fibulin C
 FT modified_site 98
 FT modified_site /label= N-linked glycosylation
 FT modified_site 535
 FT modified_site /label= N-linked glycosylation
 FT region 339
 FT region /label= N-linked glycosylation
 FT region 36.144
 FT region /label= type I motif
 FT region 36..69
 FT region /label= repeat unit 1
 FT region 112..144
 FT region /label= repeat unit 2
 FT region 144..179
 FT region /label= Glu/Asp-rich region
 FT region 180..566
 FT region /label= type II motif
 FT region 180..214
 FT region /label= repeat unit 1
 FT region 215..219
 FT region /label= consensus pentapeptide
 FT region 220..260
 FT region /label= repeat unit 2
 FT region 261..265
 FT region /label= consensus pentapeptide
 FT region 266..306
 FT region /label= repeat unit 3
 FT region 307..311
 FT region /label= consensus pentapeptide
 FT region 312..354
 FT region /label= repeat unit 4

FT region 355..359
 FT /label= consensus pentapeptide
 FT region 360..397
 FT /label= repeat unit 5
 FT region 398..402
 FT /label= consensus pentapeptide
 FT region 403..439
 FT /label= repeat unit 6
 FT region 440..444
 FT /label= consensus pentapeptide
 FT region 445..479
 FT /label= repeat unit 7
 FT region 480..484
 FT /label= consensus pentapeptide
 FT region 485..529
 FT /label= repeat unit 8
 FT region 530..535
 FT /label= consensus pentapeptide
 PN W09102755-A.
 PD 07-MAR-1991.
 PF 17-AUG-1990; U04662.
 PR 18-AUG-1989; US-395773.
 PA (LJOL-) LA JOLIA CANCER RES.
 PA (AMNA-) AMER NAT RED CROSS.
 PI Ruoslahti ET, Argraves WS;
 DR WPI; 91-087250/12.
 DR N-PSDB; Q11010.
 PT Purified fibulin, DNA encoding it and antibodies reactive with it
 PT - useful as diagnostic and therapeutic component.
 PS Claim 10; Fig 5; 56pp; English.
 CC The fibulin A, B and C forms are identical from their N-terminals
 CC to a divergence point at posn. 566 (terminal codon in fibulin A)
 CC after which they are distinct, encoding polypeptides of 566, 601
 CC and 683 residues resp. All three forms are rich in cysteine (11 %)
 CC and analysis wrt no. and spacing of the Cys residues revealed two
 CC types of repeat motif (I and II). The type I motif, CC(X)12C-
 CC (X)9-10C(X)5CC is repeated twice, separated by an imperfect form of
 CC the repeat lacking two Cys residues. The same motif is found in
 CC complement components C3a, C4a and C5a; the inverse is found in
 CC albumin, vitamin D-binding protein and alpha-fetoprotein. The
 CC disulphide-stabilized loop structure is thought to be conserved.
 CC The type II motif, related to repeats found in epidermal growth
 CC factor precursor is a 6-Cys pattern repeated nine times, although
 CC the ninth repeat in the A form is incomplete. Four of the repeats,
 CC (2, 3, 4 and 9) differ from the EGF-like motif in having a 4-6
 CC residue insertion between cysteines 4 and 5 instead of the usual
 CC single residue. Embodied within repeats 5, 6, 7 and 8 is the
 CC consensus sequence for Asp and Asn hydroxylation. The 7th repeat
 CC contains the consensus O-glycosylation sequence, CXCPX.
 CC Immediately following each repeat is a pentapeptide with the
 CC consensus sequence XD(I/V)(D/N)E. Fibulin binds to the cyto-
 CC plasmic domain of the beta1 subunit of integrin adhesion
 CC receptors in a cation-dependent, EDTA-reversible manner. It can
 CC be used to manipulate adhesion of cells to fibronectin, collagen,
 CC laminin, and possibly also other proteins. Antibodies reactive
 CC with the protein have important diagnostic and therapeutic uses.
 CC See also R1147, R1148 and R1149.
 SQ Sequence 683 AA;

Query Match 10.7%; Score 96; DB 2; Length 683;
 Best Local Similarity 57.1%; Pred. No. 5.13e+00;
 Matches 16; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

Db 10 vplllllglllaagvdadvleacc 37
 QY 6 VTLLLLLLPALLSLGVDQAQLPD-CC 32

RESULT 4
 ID P70134 standard; Protein; 440 AA.
 AC P70134;
 DT 26-APR-1991 (first entry)
 DE Natural recombinant human lecithin:cholesterol acyltransferase.

[illegible]

FH Key Location/Qualifiers
FT region 361..369
FT /note= "This line is omitted in the specification"
FT region 622..630
FT /note= "This line is omitted in the specification"
FT region 883..891
FT /note= "This line is omitted in the specification"
PN J06014783-A.
PD 25-JAN-1994.
PF 30-JUN-1992; 173155.
PR 30-JUN-1992; JP-173155.
PA (MITU) MITSUBISHI KASEI CORP.
DR WPI; 94-061478/08.
DR N-PSDB; Q56916.
PT New glutamic acid receptor and gene - for use in analysis of
PT synaptic signal translation, and diagnosis of brain disease
PS Claim 1; Page 25-31; 35pp; Japanese.
CC The sequence shows a glutamic acid receptor. The receptor is
CC useful for the analysis of nerve signal translation; within the
CC synapse, expression of synapse plasticity, nerve cell necrosis,
CC brain structure and brain disease. It can also be used in gene
CC therapy.
SQ Sequence 1239 AA;
Query Match 9.3%; Score 84; DB 9; Length 1239;
Best Local Similarity 38.5%; Pred. No. 3.93e+01;
Matches 15; Conservative 8; Mismatches 15; Indels 1; Gaps 1;
Db 957 gggrtptlarrappqparpaggrrisptcpehpatglm 995
::| |::| |::| |::| |::| |::| |::| |
QY 52 AAGILTLGKRRPGPPGLQGRLQ-RLQLQANGHAAGILTM 89

RESULT 15
ID R50056 standard; Protein; 179 AA.
AC R50056;
DT 04-MAY-1994 (first entry)
DE ICP34.5 fragment.
KW Gamma1 34.5; ICP34.5; programmed cell death; apoptosis;
KW neurovirulence; inhibition; viral replication; gene therapy;
KW neurodegenerative disease; herpes simplex virus; ss.
OS Herpes simplex virus 1.
PN W09319591-A.
PD 14-OCT-1993.
PF 26-FEB-1993; U01801.
PR 31-MAR-1992; US-861233.
PA (ARCH-) ARCH DEV CORP.
PI Chou J, Roizman B;
WPI; 93-336453/42.
Use of gamma, 34.5 gene or encoded ICP 34.5 - for preventing or
treating programmed cell death in neuronal cells or in screening
PT assays
PS Disclosure; Page 74-75; 95pp; English.
CC Programmed cell death may be treated through the use of the HSV-1
CC gene gamma1 34.5 or prod. of its expression, ICP34.5. The gene and
CC its expression have been demonstrated to be required for HSV-1
CC neurovirulence, and in partic., to act as an inhibitor of neuronal
CC programmed cell death which allows for viral replication. Use of
CC the gene therapy, or the protein itself, can be expected to result
CC in inhibition of programmed cell death in various neurodegenerative
CC diseases.
CC DNA sequences of HSV-1 strains F (Q50037-38, Q50040-42),
CC 173syn+ (Q50043-45, Q58777), MGH-10 (Q58778-83), and CVG-2 (Q58784-88)
CC and the predicted open reading frames for ICP34.5 (R50048-58)
CC are compared in Fig 1.
SQ Sequence 179 AA;
Query Match 9.2%; Score 83; DB 8; Length 179;
Best Local Similarity 32.8%; Pred. No. 4.63e+01;
Matches 19; Conservative 13; Mismatches 23; Indels 3; Gaps 3;
Db 121 hpsrpfllprlrlrvtaehla-rlrl-ragegepeppatpatpatp 176
::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |

QY 61 RRPGRP-GLQGRLLQRLQANGHAAGILTMGRAGAELEPHPCSGRGCTVTTTALAP 117
Search completed: Thu Jul 30 09:26:32 1998
Job time : 29 secs.

W P S R E H (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:28:36 1998; MasPar time 1.82 Seconds
395.942 Million cell updates/sec
ular output not generated.

Title: >US-08-938-548A-10
Description: (1-123) from US08938548A.pep
Perfect Score: 899
Sequence: 1 VPWAAVILLILLLLPPALLS.....GRGCTVTTTALPRGGSGV 123

Scoring table: PAM 150
Gap 11

Searched: 63816 seqs, 5850866 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5_COMB 2:PCT9_COMB 3:backfiles

Statistics: Mean 27.378; Variance 125.728; scale 0.218

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	88	9.8	254	1	US-08-236- Sequence 4, Applicatio	8.21e+00
2	84	9.3	238	1	US-08-453- Sequence 2, Applicatio	1.58e+01
3	84	9.3	238	1	US-08-240- Sequence 2, Applicatio	1.58e+01
4	84	9.3	1239	1	US-08-026- Sequence 3, Applicatio	1.58e+01
5	83	9.2	235	2	PCT-US94-0 Sequence 6, Applicatio	1.86e+01
6	83	9.2	235	1	US-08-243- Sequence 6, Applicatio	1.86e+01
7	83	9.2	263	2	PCT-US91-0 Sequence 2, Applicatio	1.86e+01
8	82	9.1	220	3	5175255-4 Patent No. 5175255.	2.18e+01
9	81	9.0	226	3	5498600-2 Patent No. 5498600.	2.56e+01
10	81	9.0	241	3	5175255-8 Patent No. 5175255.	2.56e+01
11	81	9.0	241	3	5175255-2 Patent No. 5175255.	2.56e+01
12	81	9.0	282	1	US-08-445- Sequence 1, Applicatio	2.56e+01
13	80	8.9	492	1	US-08-001- Sequence 4, Applicatio	3.01e+01
14	80	8.9	492	1	US-07-794- Sequence 4, Applicatio	3.01e+01
15	80	8.9	1337	2	PCT-US95-0 Sequence 2, Applicatio	3.01e+01
16	79	8.8	241	3	5194596-15 Patent No. 5194596.	3.53e+01
17	79	8.8	241	1	US-08-387- Sequence 4, Applicatio	3.53e+01
18	79	8.8	241	2	PCT-US96-0 Sequence 9, Applicatio	3.53e+01
19	78	8.7	555	1	US-08-484- Sequence 6, Applicatio	4.14e+01
20	78	8.7	1184	1	US-08-446- Sequence 20, Applicati	4.14e+01
21	78	8.7	1184	1	US-08-446- Sequence 20, Applicati	4.14e+01
22	78	8.7	1187	1	US-08-097- Sequence 13, Applicati	4.14e+01
23	78	8.7	1187	1	US-08-357- Sequence 8, Applicatio	4.14e+01

24	78	8.7	1187	2	PCT-US95-1	Sequence 8, Applicatio	4.14e+01
25	77	8.6	589	2	PCT-US91-0	Sequence 2, Applicatio	4.85e+01
26	77	8.6	589	1	US-07-668-	Sequence 2, Applicatio	4.85e+01
27	77	8.6	613	2	PCT-US94-1	Sequence 2, Applicatio	4.85e+01
28	77	8.6	2509	1	US-08-469-	Sequence 10, Applicati	4.85e+01
29	76	8.5	25	2	PCT-US94-0	Sequence 36, Applicati	5.67e+01
30	76	8.5	219	3	5310729-4	Patent No. 5310729.	5.67e+01
31	76	8.5	231	2	PCT-US95-0	Sequence 6, Applicatio	5.67e+01
32	76	8.5	231	1	US-08-243-	Sequence 2, Applicatio	5.67e+01
33	76	8.5	231	2	PCT-US94-0	Sequence 7, Applicatio	5.67e+01
34	76	8.5	231	1	US-08-220-	Sequence 2, Applicatio	5.67e+01
35	76	8.5	488	1	US-07-794-	Sequence 2, Applicatio	5.67e+01
36	76	8.5	488	1	US-08-001-	Sequence 2, Applicatio	5.67e+01
37	76	8.5	529	2	PCT-US94-0	Sequence 2, Applicatio	5.67e+01
38	76	8.5	4544	1	US-08-469-	Sequence 52, Applicati	5.67e+01
39	75	8.3	585	3	5475086-6	Patent No. 5475086.	6.64e+01
40	75	8.3	585	1	US-08-485-	Sequence 11, Applicati	6.64e+01
41	75	8.3	585	1	US-08-117-	Sequence 2, Applicatio	6.64e+01
42	75	8.3	585	1	US-08-485-	Sequence 13, Applicati	6.64e+01
43	75	8.3	689	1	US-08-059-	Sequence 5, Applicatio	6.64e+01
44	75	8.3	689	1	US-07-766-	Sequence 5, Applicatio	6.64e+01
45	75	8.3	689	2	PCT-US91-0	Sequence 5, Applicatio	6.64e+01

ALIGNMENTS

RESULT 1
ID US-08-236-918A-4 STANDARD; PRT: 254 AA.
XX
AC xxxxxx
XX
DT XX

Sequence 4, Application US/08236918A

Sequence 4, Application US/08236918A
Patent No. 5674704
GENERAL INFORMATION:
APPLICANT: Alderson, Mark R.
APPLICANT: Goodwin, Raymond G.
APPLICANT: Smith, Craig A.
TITLE OF INVENTION: Cytokine Designated 4-1BB Ligand
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kathryn A. Anderson, Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple 7.5.3
SOFTWARE: Microsoft Word, Version #6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/236,918A
FILING DATE: 06-May-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/060,843
FILING DATE: 07-May-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Anderson, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2801-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 254 amino acids

0
0
4
0
4
0
4
0
.
4
4
0
.
)
)

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

On: Thu Jul 30 09:25:01 1998; MasPar time 2.04 Seconds
77.476 Million cell updates/sec
Far output not generated.

Title: >US-08-938-548A-9
Description: (1-27) from US08938548A.pep
Perfect Score: 192
Sequence: 1 PGPPGLOGLRQLRLLQANGNHAAGILTM 27

Scoring table: PAM 150
Gap 15

Searched: 63816 seqs, 5850866 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5_COMB 2:PCIT9_COMB 3:backfiles

Statistics: Mean 20.122; Variance 80.971; scale 0.249

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	Score	Query Match	Length	DB ID	Description	Pred. No.
1	66	34.4	10	2	PCT-US91-0 Sequence 3, Applicatio	1.86e+01
2	65	33.9	716	2	PCT-US95-1 Sequence 4, Applicatio	2.26e+01
3	65	33.9	2509	1	US-08-469- Sequence 10, Applicati	2.26e+01
4	60	31.3	199	2	PCT-US93-0 Sequence 2, Applicatio	5.81e+01
5	60	31.3	199	1	US-07-941- Sequence 2, Applicatio	5.81e+01
6	60	31.3	199	1	US-07-949- Sequence 4, Applicatio	5.81e+01
7	60	31.3	199	1	US-08-115- Sequence 2, Applicatio	5.81e+01
8	60	31.3	199	1	US-08-792- Sequence 8, Applicatio	5.81e+01
9	60	31.3	199	1	US-08-017- Sequence 4, Applicatio	5.81e+01
10	60	31.3	296	1	US-08-115- Sequence 4, Applicatio	5.81e+01
11	60	31.3	296	1	PCT-US93-0 Sequence 4, Applicatio	5.81e+01
12	60	31.3	296	1	US-08-165- Sequence 14, Applicati	5.81e+01
13	60	31.3	296	1	US-07-921- Sequence 14, Applicati	5.81e+01
14	60	31.3	296	1	US-07-745- Sequence 14, Applicati	5.81e+01
15	60	31.3	296	2	PCT-US94-1 Sequence 14, Applicati	5.81e+01
16	60	31.3	296	1	US-07-941- Sequence 4, Applicatio	5.81e+01
17	58	30.2	28	1	US-07-776- Sequence 25, Applicati	8.42e+01
18	58	30.2	27	1	US-08-062- Sequence 43, Applicati	8.42e+01
19	58	30.2	27	1	US-07-924- Sequence 10, Applicati	8.42e+01
20	58	30.2	199	3	5215895-4 Patent No. 5215895	8.42e+01
21	58	30.2	199	3	5215895-3 Patent No. 5215895	8.42e+01
22	58	30.2	199	1	US-08-017- Sequence 2, Applicatio	8.42e+01
23	58	30.2	199	1	US-07-949- Sequence 2, Applicatio	8.42e+01

24	58	30.2	226	3	5498600-2	Patent No. 5498600	8.42e+01
25	58	30.2	241	3	5175255-8	Patent No. 5175255	8.42e+01
26	58	30.2	241	3	5175255-2	Patent No. 5175255	8.42e+01
27	58	30.2	282	1	US-08-445- Sequence 1, Applicatio	Sequence 1, Applicatio	8.42e+01
28	58	30.2	365	1	US-08-186- Sequence 4, Applicatio	Sequence 4, Applicatio	8.42e+01
29	58	30.2	1239	1	US-08-026- Sequence 3, Applicatio	Sequence 3, Applicatio	8.42e+01
30	58	30.2	1271	2	PCT-US94-0 Sequence 2, Applicatio	Sequence 2, Applicatio	8.42e+01
31	57	29.7	220	3	5175255-4	Patent No. 5175255	1.01e+02
32	57	29.7	829	1	US-08-220- Sequence 2, Applicatio	Sequence 2, Applicatio	1.01e+02
33	57	29.7	829	1	US-08-445- Sequence 2, Applicatio	Sequence 2, Applicatio	1.01e+02
34	57	29.7	829	1	US-08-445- Sequence 2, Applicatio	Sequence 2, Applicatio	1.01e+02
35	57	29.7	829	1	US-07-670- Sequence 2, Applicatio	Sequence 2, Applicatio	1.01e+02
36	57	29.7	971	2	PCT-US96-0 Sequence 2, Applicatio	Sequence 2, Applicatio	1.01e+02
37	57	29.7	1121	1	US-07-789- Sequence 2, Applicatio	Sequence 2, Applicatio	1.01e+02
38	57	29.7	1121	1	US-08-005- Sequence 2, Applicatio	Sequence 2, Applicatio	1.01e+02
39	57	29.7	1121	1	US-08-487- Sequence 2, Applicatio	Sequence 2, Applicatio	1.01e+02
40	57	29.7	1334	3	5476657-1	Patent No. 5476657	1.01e+02
41	57	29.7	1403	1	US-07-908- Sequence 3, Applicatio	Sequence 3, Applicatio	1.01e+02
42	56	29.2	51	1	US-08-208- Sequence 11, Applicati	Sequence 11, Applicati	1.22e+02
43	56	29.2	302	1	US-07-783- Sequence 1, Applicati	Sequence 1, Applicati	1.22e+02
44	56	29.2	405	1	US-07-688- Sequence 14, Applicati	Sequence 14, Applicati	1.22e+02
45	56	29.2	442	1	US-08-208- Sequence 2, Applicatio	Sequence 2, Applicatio	1.22e+02

ALIGNMENTS

RESULT 1
ID PCT-US91-04588-3 STANDARD: PRT: 10 AA.
XX
AC
XX
DT
XX
Sequence 3, Application PC/TUS9104588
Sequence 3, Application PC/TUS9104588
GENERAL INFORMATION:
APPLICANT: Browning, Jeffrey
APPLICANT: Ware, Carl F.
TITLE OF INVENTION: SURFACE COMPLEXED LYMPHOTOXIN
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: c/o Fish & Neave
STREET: 875 Third Avenue, 29th Floor
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022-6250
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/04588
FILING DATE: 19910627
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/544,862
FILING DATE: 27-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Haley Jr., James F.
REGISTRATION NUMBER: 27,794
REFERENCE/DOCKET NUMBER: B129CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-715-0600
TELEFAX: 212-715-0674
TELEX: 14-8367
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: AMINO ACID
TOPOLOGY: unknown

DE Sequence 4, Application US/08017522A
 XX Sequence 4, Application US/08017522A
 CC Patent No. 5371193
 CC GENERAL INFORMATION:
 CC APPLICANT: BENNETT, FRANCES K
 CC APPLICANT: PAUL, STEPHAN R
 CC APPLICANT: YANG, YU-CHUNG
 CC TITLE OF INVENTION: A MAMMALIAN CYTOKINE, IL-11
 CC NUMBER OF SEQUENCES: 4
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: LEGAL AFFAIRS - GENETICS INSTITUTE, INC.
 CC STREET: 87 CAMBRIDGE PARK DRIVE
 CC CITY: CAMBRIDGE
 CC STATE: MA
 CC COUNTRY: US
 CC ZIP: 02140
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25
 CC CURRENT APPLICATION NUMBER: US/08/017,522A
 CC FILING DATE: 19930212
 CC CLASSIFICATION: 530
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: MEINERT, M C
 CC REGISTRATION NUMBER: 31,544
 CC REFERENCE/DOCKET NUMBER: GI 5174A-DIV
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 876-1210 X8574
 CC TELEFAX: (617) 876-5851
 CC INFORMATION FOR SEQ ID NO: 4:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 199 amino acids
 CC TYPE: AMINO ACID
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 199 AA; 21429 MW; 188641 CN;
 Query Match 31.3%; Score 60; DB 1; Length 199;
 Best Local Similarity 70.0%; Pred. No. 5.81e+01;
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 Db 129 LQARLDRLLR 138
 6 LQGRRLRLQ 15
 RESULT 10
 ID US-08-115-680-4 STANDARD; PRT; 296 AA.
 XX AC xxxxxx
 XX DE
 XX DE Sequence 4, Application US/08115680
 CC Sequence 4, Application US/08115680
 CC Patent No. 5437863
 CC GENERAL INFORMATION:
 CC APPLICANT: Williams, David A.
 CC APPLICANT: Clark, Steven C.
 CC TITLE OF INVENTION: Method of Treating Cell Damage or
 CC TITLE OF INVENTION: Depletion
 CC NUMBER OF SEQUENCES: 4
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Howson and Howson
 CC STREET: Spring House Corporate Cntr, P.O. Box 457
 CC CITY: Spring House
 CC STATE: Pennsylvania

CC COUNTRY: USA
 CC ZIP: 19477
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25
 CC CURRENT APPLICATION NUMBER: US/08/115,680
 CC FILING DATE: 02-SEP-1992
 CC CLASSIFICATION: 514
 CC PRIOR APPLICATION NUMBER: US 07/941,372
 CC FILING DATE: 02-SEP-1992
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Bak, Mary E.
 CC REGISTRATION NUMBER: 31,215
 CC REFERENCE/DOCKET NUMBER: IND1Ausa
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (215) 540-9206
 CC TELEFAX: (215) 540-5818
 CC INFORMATION FOR SEQ ID NO: 4:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 296 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 296 AA; 31769 MW; 414378 CN;
 Query Match 31.3%; Score 60; DB 1; Length 296;
 Best Local Similarity 70.0%; Pred. No. 5.81e+01;
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 Db 226 LQARLDRLLR 235
 6 LQGRRLRLQ 15
 RESULT 11
 ID PCT-US93-08247-4 STANDARD; PRT; 296 AA.
 XX AC xxxxxx
 XX DE
 XX DE Sequence 4, Application PC/TUS9308247
 CC Sequence 4, Application PC/TUS9308247
 CC GENERAL INFORMATION:
 CC APPLICANT: Genetics Institute, Inc.,
 CC APPLICANT: 87 CambridgePark Drive,
 CC APPLICANT: Cambridge, MA 02140, USA
 CC TITLE OF INVENTION: Method of Treating Cell Damage or
 CC TITLE OF INVENTION: Depletion
 CC NUMBER OF SEQUENCES: 4
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Howson and Howson
 CC STREET: Spring House Corporate Cntr, P.O. Box 457
 CC CITY: Spring House
 CC STATE: Pennsylvania
 CC COUNTRY: USA
 CC ZIP: 19477
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25
 CC CURRENT APPLICATION NUMBER: PCT/US93/08247
 CC FILING DATE:
 CC CLASSIFICATION:
 CC PRIOR APPLICATION NUMBER: US 07/941,372

MPRch_pp protein - protein database search, using Smith-Waterman algorithm

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPRch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Jul 30 09:22:53 1998; Magpar time 3.90 Seconds
106.277 Million cell updates/sec

Title: >US-08-938-548A-9
Description: (1-27) from US08938548A.pap
Perfect Score: 192
Sequence: 1 FGPPGLQRLQRLQANCHAAGILTM 27

Scoring table: PAM 150
Gap 15

Searched: 124785 seqs, 15338987 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq31-2
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27

Statistics: Mean 21.806; Variance 86.670; scale 0.252

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	66	34.4	10	R20235	"p33" N-terminal (2)	5.96e+01
2	65	33.9	716	R99737	Retinoid X receptor i	7.19e+01
3	65	33.9	2509	R32881	Protein (OA-519) cros	7.19e+01
4	64	33.3	1477	R92050	KM31-7 precursor.	8.67e+01
5	64	33.3	1477	R67691	S. cerevisiae scaur2R	8.67e+01
6	64	33.3	1477	W10424	Saccharomyces cerevis	8.67e+01
7	63	32.8	878	R13793	Ecdysone receptor.	1.04e+02
8	63	32.8	878	R32889	DHR23alpha protein.	1.04e+02
9	61	31.8	32	R00579	New polypeptide based	1.51e+02
10	61	31.8	248	R06331	Human alveolar surfac	1.51e+02
11	61	31.8	248	R06394	Sequence deduced from	1.51e+02
12	61	31.8	248	R07662	35kd pulmonary surfac	1.51e+02
13	61	31.8	248	R07663	35kd pulmonary surfac	1.51e+02
14	61	31.8	248	R06065	Sequence of human alv	1.51e+02
15	61	31.8	248	R06066	Genomic sequence of h	1.51e+02
16	61	31.8	248	R06041	Plasmid pASPC-SV(10)	1.51e+02
17	61	31.8	248	R06042	Plasmid pASPCq-SV(10)	1.51e+02
18	61	31.8	248	R02980	Sequence deduced from	1.51e+02
19	61	31.8	248	R05091	vector PSP 35K-1A-10	1.51e+02

20	61	31.8	248	1	R04215	Human 32K ASP encoded	1.51e+02
21	61	31.8	248	1	R04216	Human 32K ASP encoded	1.51e+02
22	61	31.8	271	1	R04212	Human 32K alveolar su	1.51e+02
23	61	31.8	271	3	R06061	Genomic sequence of h	1.51e+02
24	61	31.8	271	1	R04217	Human 32K ASP encoded	1.51e+02
25	60	31.3	169	8	R43261	Human adipogenesis in	1.81e+02
26	60	31.3	177	8	R43262	Human adipogenesis in	1.81e+02
27	60	31.3	178	18	R02202	Human interleukin-11	1.81e+02
28	60	31.3	178	14	R75337	Human interleukin-11	1.81e+02
29	60	31.3	199	4	R24436	Sequence of a cytokin	1.81e+02
30	60	31.3	199	10	R50176	Human Interleukin-11	1.81e+02
31	60	31.3	199	8	R43260	Human adipogenesis in	1.81e+02
32	60	31.3	199	2	R12314	Human interleukin-11	1.81e+02
33	60	31.3	296	9	R45916	E.coli thioredoxin-hu	1.81e+02
34	60	31.3	296	5	R26051	Thioredoxin from PALT	1.81e+02
35	60	31.3	296	13	R75762	Thioredoxin/des-Pro-I	1.81e+02
36	60	31.3	296	10	R50177	Thioredoxin-Interleu	1.81e+02
37	60	31.3	296	5	R26213	Fusion protein of IL-	1.81e+02
38	60	31.3	296	13	R76812	Thioredoxin-IL-11 fus	1.81e+02
39	60	31.3	528	19	R05831	M. tuberculosis RNA p	1.81e+02
40	60	31.3	530	14	R76480	Virulence-associated	1.81e+02
41	59	30.7	510	9	R49835	Thermus aquaticus hea	2.18e+02
42	59	30.7	2192	23	R21732	LexA/NuMA fusion prot	2.18e+02
43	59	30.7	2272	23	R21731	GAL4/HA/NuMA fusion p	2.18e+02
44	59	30.7	3724	25	R22608	Platenolide synthase	2.18e+02
45	59	30.7	3724	25	R23718	Platenolide synthase	2.18e+02

ALIGNMENTS

RESULT 1
ID R20235 standard; Protein; 10 AA.
AC R20235;
DT 24-APR-1992 (first entry)
DE "p33" N-terminal (2).
KW MAP; lymphocyte; IL-2; CTL; polymorphism.
OS Homo sapiens. Location/Qualifiers
FH Key Misc_difference 6 /note= "polymorphism - see CC"
FT W09200329-A.
PN 09-JAN-1992.
PD 27-JUN-1991; U04588.
PF 27-JUN-1990; US-544862.
PR (BIOG-) BIOGEN INC.
PA (REGC) UNIV OF CALIFORNIA.
PI Browning J, Ware CF;
DR WPI: 92-041521/05.
PT New membrane associated protein, p33 - forms complex with
PT lymphotoxin, useful as antiinflammatory agent, tumour growth
PT inhibitor, T-cell inhibitor or activator
PS Claim 1; Page 60; 75pp; English.
CC A protein, designated p33, is found on the surface of several types
CC of lymphocyte cells, including OKT3-stimulated primary T-cells,
CC antigen-specific IL-2 dependent CTL clones, and a PMA-stimulated
CC human T-cell hybridoma, II-23.D7. It forms a novel complex with
CC lymphotoxin (LFT). The N-terminal of the p33 protein contains the
CC sequence represented in R20234 or R20235. The sequence at the 6th
CC cycle appeared to be a mixt. of both G and L indicating possible
CC polymorphism. The p33 protein may have one of these sequences or
CC both.
SQ Sequence 10 AA;

Query Match 34.4%; Score 66; DB 4; Length 10;
Best Local Similarity 88.9%; Pred. No. 5.96e+01;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 glegrlqrl 9
|||
QY 5 GLQGLRLQRL 13

RESULT 2

P70662;
 DE 29-APR-1991 (first entry)
 DT 35kd pulmonary surfactant protein.
 KW Hyaline membrane disease; respiratory distress syndrome; RDS.
 OS Homo sapiens.
 PN W08702037-A.
 PD 09-APR-1987.
 PE 26-SEP-1986; U02034.
 PF 26-SEP-1985; US-781130.
 PR 15-AUG-1986; US-897183.
 PA (GENE-) Genetics Institute Inc.
 PI (BRIG-) Brigham and Women's Hospital.
 PI Tausch HW, Cobs KA, Steinbrink DR, Floros J, Phelps DS;
 DR WPI: 87-108682/15.
 DR N-PSDB: N71009.
 PT Pulmonary surfactant proteins - used for treating Hyaline Membrane
 PT Disease or Respiratory Distress Syndrome.
 PS Claim 1; Page 33-34; 50pp; English.
 CC Gene product may be used in treatment of Hyaline Membrane Disease
 CC and Respiratory Distress Syndrome (RDS) in both premature infants
 CC and adults eg. cardio-pulmonary operations. The protein products
 CC may also be used to raise diagnostic antibodies.
 SQ Sequence 248 AA;
 Query Match 31.8%; Score 61; DB 2; Length 248;
 Best Local Similarity 60.0%; Pred. No. 1.51e+02;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Db 95 gppglpahldeeqa 109
 ||||| :|||
 QY 2 GPPGLQGRQLRLQA 16

RESULT 13
 ID P70663 standard; Protein; 248 AA.
 AC P70663;
 DT 29-APR-1991 (first entry)
 DE 35kd pulmonary surfactant protein.
 KW Hyaline membrane disease; respiratory distress syndrome; RDS.
 OS Homo sapiens.
 PN W08702037-A.
 PD 09-APR-1987.
 PE 26-SEP-1986; U02034.
 PF 26-SEP-1985; US-781130.
 PR 15-AUG-1986; US-897183.
 PA (GENE-) Genetics Institute Inc.
 PI (BRIG-) Brigham and Women's Hospital.
 PI Tausch HW, Cobs KA, Steinbrink DR, Floros J, Phelps DS;
 DR WPI: 87-108682/15.
 DR N-PSDB: N71010.
 PT Pulmonary surfactant proteins - used for treating Hyaline Membrane
 PT Disease or Respiratory Distress Syndrome.
 PS Claim 1; Page 34A-B; 50pp; English.
 CC Gene product may be used in treatment of Hyaline Membrane Disease
 CC and Respiratory Distress Syndrome (RDS) in both premature infants
 CC and adults eg. cardio-pulmonary operations. The protein products
 CC may also be used to raise diagnostic antibodies.
 SQ Sequence 248 AA;
 Query Match 31.8%; Score 61; DB 2; Length 248;
 Best Local Similarity 60.0%; Pred. No. 1.51e+02;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Db 95 gppglpahldeeqa 109
 ||||| :|||
 QY 2 GPPGLQGRQLRLQA 16

RESULT 14
 ID P60665 standard; Protein; 248 AA.
 AC P60665;
 DT 31-JUL-1991 (first entry)
 DE Sequence of human alveolar surfactant protein (hASP)

DE on pMT(E):HS and PASPC-SV(10)
 KW Regulatable expression system.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT peptide 1..20
 FT protein 21..248
 PN W08604920-A.
 PD 28-AUG-1986.
 PF 11-FEB-1986; U00296.
 PR 13-FEB-1985; US-701296.
 PR 25-NOV-1985; US-801674.
 PA (BIOT-) BIOTECHN RES PARTNE.
 PA (KUSH/) KUSHNER P J.
 PI Kushner PJ, Cofer CL, Friedman J, Talmadge KD;
 DR WPI: 86-23888/36.
 DR N-PSDB: N60571.
 PT Regulatable expression systems - contg. human metallo:thionein-II
 PT promoter
 PS Example; Fig 5; 94pp; English.
 CC A regulatable expression system for a coding sequence is claimed.
 CC The system can process genomic as well as intronless DNA.
 SQ Sequence 248 AA;
 Query Match 31.8%; Score 61; DB 3; Length 248;
 Best Local Similarity 60.0%; Pred. No. 1.51e+02;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Db 95 gppglpahldeeqa 109
 ||||| :|||
 QY 2 GPPGLQGRQLRLQA 16

RESULT 15
 ID P60666 standard; Protein; 248 AA.
 AC P60666;
 DT 31-JUL-1991 (first entry)
 DE Genomic sequence of human alveolar surfactant protein (hASP)
 DE encoded by genomic DNA, used to obtain PASPCg-SV(10)
 KW Regulatable expression system.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT peptide 1..20
 FT protein 21..248
 PN W08604920-A.
 PD 28-AUG-1986.
 PF 11-FEB-1986; U00296.
 PR 13-FEB-1985; US-701296.
 PR 25-NOV-1985; US-801674.
 PA (BIOT-) BIOTECHN RES PARTNE.
 PA (KUSH/) KUSHNER P J.
 PI Kushner PJ, Cofer CL, Friedman J, Talmadge KD;
 DR WPI: 86-23888/36.
 DR N-PSDB: N60572.
 PT Regulatable expression systems - contg. human metallo:thionein-II
 PT promoter
 PS Example; Fig 6; 94pp; English.
 CC A regulatable expression system for a coding sequence is claimed.
 CC The system can process genomic as well as intronless DNA.
 SQ Sequence 248 AA;
 Query Match 31.8%; Score 61; DB 3; Length 248;
 Best Local Similarity 60.0%; Pred. No. 1.51e+02;
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Db 95 gppglpahldeeqa 109
 ||||| :|||
 QY 2 GPPGLQGRQLRLQA 16

Search completed: Thu Jul 30 09:23:11 1998
 Job time : 18 secs.

W P S R L H

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:20:43 1998; MasPar time 3.05 Seconds
271.021 Million cell updates/sec
Full output not generated.

Title: >US-08-938-548A-8
Description: (1-33) from US08938548A.pap
Perfect Score: 256
Sequence: 1 QPLPCCRQKTCSCRLYELLHGAGNHAAGILTL 33

Scoring table: PAM 150
Gap.11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 31.693; Variance 44.992; scale 0.704

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	75	29.3	373	1	ICEB_MOUSE CASPASE-11 PRECURSOR (5.12e-02
2	73	28.5	260	1	NMA_HUMAN PUTATIVE TRANSMEMBRANE	1.20e-01
3	71	27.3	450	1	PPAW_CAEEL PUTATIVE ACID PHOSPHAT	2.76e-01
4	70	27.3	214	1	R10A_TRYER 50S RIBOSOMAL PROTEIN	4.17e-01
5	69	27.0	197	1	MCS_MOUSE SPERM MITOCHONDRIAL CA	6.26e-01
6	69	27.0	461	1	YUL2_CAEEL PUTATIVE FORKHEAD-RELA	6.26e-01
7	67	26.2	299	1	ALC_RABIT IG ALPHA CHAIN C REGIO	1.40e+00
8	67	26.2	923	1	RE11_SCHPO MEIOTIC RECOMBINATION	1.40e+00
9	66	25.8	347	1	GALE_RAT UDP-GLUCOSE 4-EPIMERAS	2.07e+00
10	66	25.8	348	1	GALE_HUMAN UDP-GLUCOSE 4-EPIMERAS	2.07e+00
11	66	25.8	360	1	PURK_PSEAE PHOSPHORIBOSYLAMINOIM	2.07e+00
12	66	25.8	425	1	IF15_MOUSE INTERFERON-ACTIVABLE	2.07e+00
13	65	25.4	155	1	Y115_MERJA HYPOTHETICAL PROTEIN M	3.06e+00
14	65	25.4	273	1	MD12_SCHPO MITOCHONDRIAL INHERIT	3.06e+00
15	64	25.0	75	1	TXO3_AGEAP OMEGA-AGATOXIN I1IA	4.51e+00
16	64	25.0	83	1	TX4B_AGEAP OMEGA-AGATOXIN IVE PRE	4.51e+00
17	64	25.0	103	1	BOLA_HAEIN BOLA PROTEIN HOMOLOG.	4.51e+00
18	64	25.0	125	1	AGSW_VULVU AGOUTI SWITCH PROTEIN	4.51e+00
19	64	25.0	131	1	AGSW_MOUSE AGOUTI SWITCH PROTEIN	4.51e+00
20	64	25.0	132	1	AGSW_HUMAN AGOUTI SWITCH PROTEIN	4.51e+00
21	64	25.0	493	1	VPE_VICSA VACUOLAR PROCESSING EN	4.51e+00
22	63	24.6	236	1	ECSC_BACSU PROTEIN ECSC.	6.60e+00
23	63	24.6	477	1	ANGT_MOUSE ANGIOTENSINOGEN PRECUR	6.60e+00

24	63	24.6	639	1	TET9_ENTFA TETRACYCLINE RESISTANC	6.60e+00
25	62	24.2	612	1	UNC6_CAEEL UNC-6 PROTEIN PRECURSO	9.61e+00
26	62	24.2	1464	1	NME1_RAT GLUTAMATE (NMDA) RECEP	9.61e+00
27	62	24.2	1464	1	NME1_MOUSE GLUTAMATE (NMDA) RECEP	9.61e+00
28	62	24.2	1466	1	SPA2_YEAST SPA2 PROTEIN.	9.61e+00
29	62	24.2	1976	1	MYSO_HUMAN MYOSIN HEAVY CHAIN, NO	9.61e+00
30	61	23.8	160	1	VG34_HSVEB GENE 34 PROTEIN.	1.39e+01
31	61	23.8	270	1	URED_KLEPN UREASE OPERON URED PRO	1.39e+01
32	61	23.8	270	1	URED_KLEPN UREASE OPERON URED PRO	1.39e+01
33	61	23.8	406	1	Y129_HUMAN HYPOTHETICAL PROTEIN K	1.39e+01
34	61	23.8	442	1	TBB_TRYBR TUBULIN BETA CHAIN.	1.39e+01
35	61	23.8	483	1	PRPD_SALTY PRPD PROTEIN.	1.39e+01
36	61	23.8	595	1	BETP_CORGL GLYCINE BETAINE TRANSP	1.39e+01
37	61	23.8	1169	1	RAD5_YEAST DNA REPAIR PROTEIN RAD	2.01e+01
38	60	23.4	334	1	YHO5_YEAST HYPOTHETICAL 37.9 KD P	2.01e+01
39	60	23.4	371	1	NTF6_TOBAC MITOGEN-ACTIVATED PROT	2.01e+01
40	60	23.4	471	1	YMT1_CAEEL HYPOTHETICAL 51.7 KD P	2.01e+01
41	60	23.4	639	1	YTM1_UREUR TETRACYCLINE RESISTANC	2.01e+01
42	60	23.4	1039	1	YAF3_SCHPO HYPOTHETICAL 118.6 KD	2.01e+01
43	60	23.4	1122	1	DPOL_ADE07 DNA POLYMERASE (EC 2.7	2.01e+01
44	60	23.4	1150	1	YIC6_YEAST HYPOTHETICAL 133.0 KD	2.01e+01
45	60	23.4	2769	1	THYG_BOVIN THYROGLOBULIN PRECURSO	2.01e+01

ALIGNMENTS

RESULT	ID	ICEB_MOUSE	STANDARD;	PRT;	373 AA.
AC	P70343;	O08735;			
DT	01-NOV-1997	(REL. 35, CREATED)			
DT	01-NOV-1997	(REL. 35, LAST SEQUENCE UPDATE)			
DT	01-NOV-1997	(REL. 35, LAST ANNOTATION UPDATE)			
DE	CASPASE-11 PRECURSOR (EC 3.4.22.-)	(ICH-3 PROTEASE).			
GN	CASP11 OR ICH3 OR CASP1.				
OS	MUS MUSCULUS (MOUSE).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; RODENTIA.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-C57BL/6 X CBA; TISSUE-THYMUS;				
RX	MEDLINE; 96355393.				
RA	WANG S., MIURA M., JUNG Y.K., ZHU H., GAGLIARDINI V., SHI L.,				
RA	GREENBERG A.H., YUAN J.				
RL	J. BIOL. CHEM. 271:20580-20587(1996).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-C3H/AN;				
RX	MEDLINE; 97190206.				
RA	VAN DE CRAEN M., VANDENABEELE P., DECLERCO W., VAN DEN BRANDE I.,				
RA	VAN LOO G., MOLEMAN F., SCHOTTE P., VAN CRIEKNINGE W., BEYAERT R.,				
RA	FIERS W.				
RL	FEBS LETT. 403:61-69(1997).				
CC	-1- FUNCTION: INVOLVED IN THE ACTIVATION CASCADE OF CASPASES				
CC	RESPONSIBLE FOR APOPTOSIS EXECUTION (BY SIMILARITY).				
CC	-1- SUBUNIT: HETERODIMER OF TWO SUBUNITS (POTENTIAL).				
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C14; ALSO KNOWN AS THE				
CC	CASPASE FAMILY.				
DR	EMBL; U59463; G1575318; -				
DR	EMBL; Y13089; E315510; -				
DR	MGD; MGI:107700; CASPL.				
DR	PROSITE; PS01121; CASPASE_HIS; 1.				
DR	PROSITE; PS01122; CASPASE_CIS; 1.				
KW	HYDROLASE; THIOL PROTEASE; APOPTOSIS; ZYMOGEN.				
FT	PROPEP 1 79				
FT	CHAIN 80 266				
FT	PROPEP 267 285				
FT	CHAIN 286 373				
FT	ACT_SITE 206 206				
FT	ACT_SITE 254 254				
FT	CONFLICT 152 152				
SQ	SEQUENCE 373 AA; 427556 MW; 7DDB4A63 CRC32;				

Query Match 29.3%; Score 75; DB 1; Length 373;


```
RESULT 13
ID Y115.MEJJA STANDARD; PRT; 155 AA.
AC Q57579;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE HYPOTHETICAL PROTEIN MJ0115.
GN MJ0115.
OS METHANOCOCCUS JANNASCHII.
OC ARCHAEABACTERIA; EURYARCHAEOTA; METHANOCOCCALES; METHANOCOCCACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96337999.
RA BULT C.J., WHITE O., OLSEN G.J., ZHOU L., FLEISCHMANN R.D.,
RA SUTTON G.G., BLAKE J.A., FITZGERALD L.M., CLAYTON R.A., GOCAYNE J.D.,
RA KERLAVAGE A.R., DOUGHERTY B.A., TOMB J.-F., ADAMS M.D., REICH C.I.,
RA OVERBECK R., KIRKNESS E.F., WEINSTOCK K.G., MERRICK J.M., GLODEK A.,
RA SCOTT J.L., GEOGHAGEN N.S.M., WEIDMAN J.F., FUHRMANN J.L., NGUYEN D.,
RA UTTERBACK T.R., KELLEY J.M., PETERSON J.D., SADOW P.W., HANNA M.C.,
RA COTTON M.D., ROBERTS K.M., HURST M.A., KAINE B.P., BORODOVSKY M.,
RA KLENK H.-P., FRASER C.M., SMITH H.O., WOESE C.R., VENTER J.C.;
RL SCIENCE 273:1058-1073(1996).
DR EMBL; U67469; G1498882;
DR TIGR; MJ0115;
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 155 AA; 17727 MW; 9D9D61E8 CRC32;

Query Match 25.4%; Score 65; DB 1; Length 155;
Best Local Similarity 53.3%; Pred. No. 3.06e+00;
Matches 8; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

Db 72 CCKITPCPYDYEL 86
QY 6 CCR-QKTCSCRYEL 19
||: ||: |||
||: ||: |||

RESULT 14
ID MD12.SCHPO STANDARD; PRT; 273 AA.
AC Q92377;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE MITOCHONDRIAL INHERITANCE COMPONENT MD12.
GN MD12.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-972;
RA BERGER K.H., SOGO L.F., YAFFE M.P.;
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- SIMILARITY: TO YEAST MD12.
DR EMBL; U64674; G1655884;
KW MITOCHONDRION.
SQ SEQUENCE 273 AA; 30517 MW; 9024B3CC CRC32;

Query Match 25.4%; Score 65; DB 1; Length 273;
Best Local Similarity 53.8%; Pred. No. 3.06e+00;
Matches 7; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Db 213 LLHGTGEHASSVI 225
QY 19 LLHGAGNHAAGIL 31
||||:||||:

RESULT 15
ID TX03.AGEAP STANDARD; PRT; 76 AA.
AC P33034;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE OMEGA-AGATOXIN IIIA.
```

```
OS AGELENOPSIS APERTA (FUNNEL-WEB SPIDER).
OC EUKARYOTA; METAZOA; ARTHROPODA; CHELICERATA; ARACHNIDA; ARANEAE.
RN [1]
RP SEQUENCE.
RX TISSUE-VENOM;
RA VENEMA V.J., SWIDEREK K.M., LEE T.D., HATHAWAY G.M., ADAMS M.E.;
RL J. BIOL. CHEM. 267:2610-2615(1992).
CC -!- FUNCTION: OMEGA-AGATOXIN ARE ANTAGONIST OF VOLTAGE-SENSITIVE
CC CALCIUM CHANNELS. THEY BLOCK INSECT NEUROMUSCULAR TRANSMISSION
CC PRESYNAPTICALLY. POTENT BLOCKER OF N- AND L-TYPE CALCIUM CHANNELS.
CC -!- PTM: SIX DISULFIDE BONDS ARE PRESENT (PROBABLE).
DR PIR; A42335; A42335.
KW VENOM; NEUROTOXIN; CALCIUM CHANNEL INHIBITOR; PRESYNAPTIC NEUROTOXIN.
SQ SEQUENCE 76 AA; 8518 MW; C7EA0E12 CRC32;

Query Match 25.0%; Score 64; DB 1; Length 76;
Best Local Similarity 38.1%; Pred. No. 4.51e+00;
Matches 8; Conservative 6; Mismatches 5; Indels 2; Gaps 2;

Db 17 OCCRRNGYCSCYSLFGLKSG 37
QY 5 DCCRQKT-CSC-RLYELLHGA 23
||||: ||| | :|
|||: ||| | :|
```

Search completed: Thu Jul 30 09:20:53 1998
Job time : 10 secs.

MPARCH_PP protein - protein database search, using Smith-Waterman algorithm

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPARCH_PP protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Jul 30 09:21:12 1998; MasPar time 5.80 Seconds
239.700 Million cell updates/sec
Linear output not generated.

Title: >US-08-938-548A-8
Description: (1-33) from US08938548A.pep
Perfect Score: 256
Sequence: 1 QLPDCCRQKTCSCRLYELHAGNHAAGIITL 33

Scoring table: PAM 150
Gap 11

Searched: 140542 seqs, 42109429 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: spiremb1
1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phage 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified

Statistics: Mean 30.295; Variance 46.908; scale 0.646

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES					Pred. No.	
Result No.	Score	Query Match	Length	ID	Description	
1	77	30.1	363	4	Q02839	PORCINE MEMBRANE COFAC
2	74	28.9	750	3	Q93473	FL4B4.1 (FRAGMENT).
3	73	28.5	683	3	Q22187	T05A1.3.
4	73	28.5	1876	3	Q24453	PHOSPHOINOSITIDE 3-KIN
5	73	28.5	1876	3	Q01938	PHOSPHOINOSITIDE 3-KIN
6	73	28.5	1876	3	Q24209	PHOSPHOINOSITIDE 3-KIN
7	72	28.1	147	10	Q61639	3' ORF.
8	72	28.1	815	9	Q33367	DNA GYRASE B SUBUNIT.
9	70	27.3	426	3	Q01969	SIMILAR TO GALACTOKINA
10	70	27.3	883	12	Q91493	DYSTROPHIN (FRAGMENT).
11	68	26.6	641	9	Q48791	TET.
12	68	26.6	645	9	Q48712	TETRACYCLINE RESISTANC
13	68	26.6	1382	12	Q90975	TYROSINE KINASE.
14	66	25.8	479	9	P94426	HOMOLOGUE OF REGULATOR
15	65	25.4	197	4	Q28584	KAP5.5 KERATIN PROTEIN
16	65	25.4	589	1	Q13388	BETA-D-FRUCTOFURANOSID
17	64	25.0	310	3	Q01473	COSMID C04E6.
18	64	25.0	369	9	P75863	FROM BASES 956879 TO 1
19	64	25.0	397	9	Q44258	1-CARBOXY-3-CHLORO-3,4
20	64	25.0	398	3	O18373	SELD PROTEIN.

21	64	25.0	398	3	O18597	SELENOPHOSPHATE SYNTH
22	63	24.6	146	9	O05606	REPRESSOR/INDUCER PROT
23	63	24.6	188	3	Q18238	COSMID C27A2.
24	63	24.6	227	8	O04393	RIBONUCLEASE.
25	63	24.6	366	3	Q22627	F21B10.6.
26	63	24.6	405	3	O19671	F21C3.1.
27	63	24.6	466	8	Q39135	AMINO ACID TRANSPORTER
28	63	24.6	639	9	O53770	TETRACYCLINE RESISTANC
29	63	24.6	639	11	Q67709	TETRACYCLINE RESISTANC
30	63	24.6	639	9	O57224	ORF11.
31	63	24.6	639	9	Q47810	TETM GENE.
32	63	24.6	897	3	O17336	LET 858.
33	63	24.6	2946	3	O18857	SIMILARITY TO EGF-LIKE
34	62	24.2	99	4	Q27991	NONMUSCLE MYOSIN HEAVY
35	62	24.2	106	2	Q12989	NONMUSCLE MYOSIN HEAVY
36	62	24.2	109	4	Q27990	NONMUSCLE MYOSIN HEAVY
37	62	24.2	157	9	Q33700	DNA FOR DNAAJ, COMPLETE
38	62	24.2	272	10	Q62707	NONMUSCLE MYOSIN HEAVY
39	62	24.2	282	10	Q62706	NONMUSCLE MYOSIN HEAVY
40	62	24.2	304	3	Q22732	T24D5.1.
41	62	24.2	1464	2	Q12879	N-METHYL-D-ASPARTATE R
42	62	24.2	1464	10	Q63728	N-METHYL-D-ASPARTATE R
43	62	24.2	1464	10	O08948	N-METHYL-D-ASPARTATE R
44	61	23.8	282	12	Q91373	TRKB (FRAGMENT).
45	61	23.8	2007	12	Q02015	MYOSIN HEAVY CHAIN, NO

ALIGNMENTS

RESULT	1	PRELIMINARY;	PRT;	363 AA.
ID	O02839			
AC	O02839			
DT	01-JUL-1997 (TREMBLREL. 04, CREATED)			
DT	01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)			
DT	01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)			
DE	PORCINE MEMBRANE COFACTOR PROTEIN.			
OS	SUS SCROFA (PIC).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; ARTIODACTYLA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE: 97343414.			
RA	TOYOMURA K., FUJIMURA T., MURAKAMI H., NATSUME T., SHIGEHISA T.,			
RA	INOUE N., TAKEIDA J., KINOSHITA T.;			
RL	INT. IMMUNOL. 9:869-876(1997).			
DR	EMBL: D70897; G1018989; -.			
KW	MEMBRANE.			
SQ	SEQUENCE 363 AA; 39692 MW; AD14F57A CRC32;			

Query Match 30.1%; Score 77; DB 4; Length 363;
Best Local Similarity 39.3%; Pred. No. 5.22e-02;
Matches 11; Conservative 7; Mismatches 9; Indels 1; Gaps 1;

Db	98	PLQACRCKACS-NLPDPLNGQVSPNG	124
QY	2	PLPDCCRQKTCSCRLYELHAGNHAAG	29

RESULT	2	PRELIMINARY;	PRT;	750 AA.
ID	Q93473			
AC	Q93473			
DT	01-FEB-1997 (TREMBLREL. 02, CREATED)			
DT	01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)			
DT	01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)			
DE	FL4B4.1 (FRAGMENT).			
OS	CAENORHABDITIS ELEGANS.			
OC	EUKARYOTA; METAZOA; ACOELOMATES; NEMATODA; SECERNENTEA; RHABDITIDA.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	SIMS M.;			
RL	SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.			
RN	[2]			
RP	SEQUENCE FROM N.A.			

Db 691 PSCRRRSTWSCRLCAEHLRGAG 712
QV . 4 PDCCRQKTCSCRLY-ELLHGAG 24

Db 691 PSCRRRSTWSCRLCAEHLRGAG 712
QV . 4 PDCCRQKTCSCRLY-ELLHGAG 24

```
DR PROSITE; PS01159;
FT NON_TER 1
```

DR	PROSITE; PS01159; WW_DOMAIN_1; 1.
FT	NON_TER 1 1

RC STRAIN-168;
 RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; D50453; G1805458; -;
 DR EMBL; Z99106; E1182356; -;
 SQ SEQUENCE 479 AA; 55166 MW; 3C2D1F5A CRC32;

Query Match 25.8%; Score 66; DB 9; Length 479;
 Best Local Similarity 41.2%; Pred. No. 3.99e+00;
 Matches 7; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 207 RMYQILKNAGKQVKTIM 223
 I:|:| |::| |:
 QY 15 RLYELLHCAGNHAGIL 31

RESULT 15
 ID Q28584 PRELIMINARY; PRT; 197 AA.
 AC Q28584;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 RL 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DE 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
 DE KAP5.5 KERATIN PROTEIN (FRAGMENT).
 GN KRTAP5.5.
 OS OVIS ARIES (SHEEP).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; ARTIODACTYLA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-WOOL FOLLICLE;
 EX MEDLINE; 94358466.
 RA JENKINS B.J., POWELL B.C.;
 RL J. INVEST. DERMATOL. 103:310-317(1994).
 DR EMBL; X73435; G313722; -;
 KW KERATIN.
 FT NON TER
 SQ SEQUENCE 197 AA; 17474 MW; 731C19CA CRC32;

Query Match 25.4%; Score 65; DB 4; Length 197;
 Best Local Similarity 58.3%; Pred. No. 5.80e+00;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 87 VPVCCRVPCSC 98
 :| | | :| | |
 QY 3 LPDCCRQKTCSC 14

Search completed: Thu Jul 30 09:21:29 1998
 time : 17 secs.

(TM)

Result No.	Query		Score	Match	Length	DB	ID	Description	Pred. No.
	1	2							
1	72	28.1	147	2	S24303	hypothetical protein	9.01e-01		
2	70	27.3	870	2	A41130	dystrophin homolog -	1.84e+00		
3	69	27.0	93	2	S60790	oct2 protein isoform	2.63e+00		
4	69	27.0	143	2	A37199	sperm mitochondrial c	2.63e+00		
5	68	26.6	641	2	JN0800	tetracycline-minocycl	3.73e+00		
6	67	26.2	299	1	AHRB	Ig alpha chain C regi	5.28e+00		
7	67	26.2	338	2	S09276	Ig alpha chain C regi	5.28e+00		
8	66	25.8	347	2	S11223	UDPglucose 4-epimeras	7.44e+00		
9	66	25.8	425	2	S156329	gene D3 protein - mou	7.44e+00		
10	66	25.8	479	2	B69764	transcriptional regul	7.44e+00		
11	65	25.4	155	2	C64314	hypothetical protein	1.05e+01		
12	65	25.4	197	2	I46413	keratin KAPS.5 - shree	1.05e+01		
13	65	25.4	589	2	S33920	beta-fructofuranosida	1.05e+01		
14	64	25.0	35	5	1OMB	Omega-aga-ivb (nmr, m	1.46e+01		
15	64	25.0	48	5	1OMB	Omega-aga-ivb (nmr, 2	1.46e+01		
16	64	25.0	48	5	1AGG	omega-agatoxin-ivb -	1.46e+01		
17	64	25.0	48	1	A44664	omega-agatoxin IVB -	1.46e+01		
18	64	25.0	76	2	B54252	omega-agatoxin III, 8	1.46e+01		
19	64	25.0	76	2	A42335	omega-agatoxin IIIA -	1.46e+01		
20	64	25.0	76	2	A54252	omega-agatoxin III, 8	1.46e+01		
21	64	25.0	103	2	B64052	cell division protein	1.46e+01		
22	64	25.0	131	2	A46298	pigment deposition co	1.46e+01		
23	64	25.0	132	2	I37143	acouti protein precur	1.46e+01		

stabilized by interchain disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into larger oligomers.

CLASSIFICATION #superfamily immunoglobulin C region; immunoglobulin homology duplication; glycoprotein; heterotrimer; immunoglobulin; plasma

FEATURE 86-152 #domain immunoglobulin homology #label IGG1
189-261 #domain immunoglobulin homology #label IGG2
38,286 #binding_site carbohydrate (Asn) (covalent) #status predicted

SUMMARY #length 299 #checksum 2361

Query Match 26.2%; Score 67; DB 1; Length 299;
Best Local Similarity 50.0%; Pred. No. 5.28e+00;
Matches 7; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Db 44 PFPDCCPANSCTC 57

QY 2 PLPDCCRQKTC-SC 14

RESULT 7

ENTRY #type fragment

TITLE Ig alpha chain C region - rabbit (fragment)

ORGANISM #formal_name Oryctolagus cuniculus #common_name domestic rabbit

DATE 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change

ACCESSIONS S09276

REFERENCE S09276

#authors Burnett, R.C.; Hanly, W.C.; Zhai, S.K.; Knight, K.L.

#journal EMBO J. (1989) 8:4041-4047

#title The IgA heavy-chain gene family in rabbit: cloning and sequence analysis of 13 C-alpha genes.

#cross-references MUID:90076124

#accession S09276

#status not compared with conceptual translation

#molecule_type DNA

#residues 1-338 #label BUR

CLASSIFICATION #superfamily immunoglobulin C region; immunoglobulin homology immunoglobulin

FEATURE 228-300

SUMMARY #domain immunoglobulin homology #label IGG2

#length 338 #checksum 2169

Query Match 26.2%; Score 67; DB 2; Length 338;

Best Local Similarity 50.0%; Pred. No. 5.28e+00;

Matches 7; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Db 83 PFPDCCPANSCTC 96

QY 2 PLPDCCRQKTC-SC 14

RESULT 8

ENTRY #type complete

TITLE UDPglucose 4-epimerase (EC 5.1.3.2) - rat

ALTERNATE_NAMES UDPgalactose 4-epimerase

ORGANISM #formal_name Rattus norvegicus #common_name Norway rat

DATE 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change

ACCESSIONS S11223

REFERENCE S11223

#authors Zeschmigg, M.; von Wilcken-Bergmann, B.; Starzinski-Powitz, A.

#journal Nucleic Acids Res. (1990) 18:5289

#title cDNA from rat cells with reconstitutive galactose-epimerase activity in E. coli.

#cross-references MUID:90384840

#accession S11223

#status preliminary

#molecule_type mRNA

#residues 1-347 #label ZPS

#cross-references EMBL:X53949; NID:g57791; PID:g57792

GENETICS

#gene gale

CLASSIFICATION #superfamily Escherichia coli UDPglucose 4-epimerase;

KEYWORDS UDPglucose 4-epimerase homology

FEATURE galactose metabolism; isomerase

SUMMARY #domain UDPglucose 4-epimerase homology #label UDP

5-343 #length 347 #molecular-weight 38225 #checksum 6947

Query Match 25.8%; Score 66; DB 2; Length 347;

Best Local Similarity 53.8%; Pred. No. 7.44e+00;

Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 261 CGCRIYNGTGTG 273

QY 12 CSRLYELHAG 24

RESULT 9

ENTRY #type complete

TITLE gene D3 protein - mouse

ORGANISM #formal_name Mus sp. #common_name mouse

DATE 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change

ACCESSIONS I56329

REFERENCE I56329

#authors Tannenbaum, C.S.; Major, J.; Ohmori, Y.; Hamilton, T.A.

#journal J. Leukoc. Biol. (1993) 53:563-568

#title A lipopolysaccharide-inducible macrophage gene (D3) is a new member of an interferon-inducible gene cluster and is selectively expressed in mononuclear phagocytes.

#cross-references MUID:93274206

#accession I56329

#status preliminary; translated from GB/EMBL/DDBJ

#molecule_type mRNA

#residues 1-425 #label RES

#cross-references GB:S62227; NID:g385702; PID:g385703

GENETICS

#gene D3

SUMMARY #length 425 #molecular-weight 47046 #checksum 8487

Query Match 25.8%; Score 66; DB 2; Length 425;

Best Local Similarity 31.0%; Pred. No. 7.44e+00;

Matches 9; Conservative 10; Mismatches 9; Indels 1; Gaps 1;

Db 178 PCCCEPTAMCO-SPILHSSSSASSNLS 205

QY 4 PDCRQKTCSCRLYELHAGNHAAGILT 32

RESULT 10

ENTRY #type complete

TITLE transcriptional regulator (GntR family) / homolog ycnF -

ORGANISM Bacillus subtilis

DATE 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change

ACCESSIONS B69764

REFERENCE A69380

#authors Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;

Alloni, G.; Azevedo, V.; Bertero, M.G.; Bessieres, P.;

Boletin, A.; Borchert, S.; Boriss, R.; Boursier, L.; Brans,

A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;

Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;

Choi, S.R.; Codani, J.J.; Connerton, I.F.; Cummings, N.J.;

Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoeft, A.;

Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.;

Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita,

M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim,

S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandl, G.;

Guiseppi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,

1-23 #domain signal sequence #status experimental #label SIG\

24-589 #product beta-fructofuranosidase #status experimental

#label MAT\

36,42,170,188,211,

254,259,318,322,

388,463,518,527

#binding_site carbohydrate (Asn) (covalent) #status

70-92 predicted\

#binding_site phosphate (Thr) (covalent) #status

458,475,490 predicted\

#binding_site phosphate (Ser) (covalent) #status

predicted

SUMMARY #length 589 #molecular-weight 63650 #checksum 1032

Query Match 25.48; Score 65; DB 2; Length 589;

Best Local Similarity 42.1%; Pred.No. 1.05e+01;

Matches 8; Conservative 6; Mismatches 4; Indels 1; Gaps 1;

Db 548 RLFDVLNG-GEQAIETLDL 565

QV 15 RLYELLHGAGNHAAGILT 33

||||:|:|:|

RESULT 14

ENTRY

TITLE

ORGANISM

REFERENCE

#authors

#submitters

#cross-references PDB:1OMB

COMMENT

COMMENT

FEATURE

7-9,31-33,23-24 #region beta sheet\

10-13 #region turn (type II)\

1-17 #disulfide_bonds\

9-22 #disulfide_bonds\

16-33 #disulfide_bonds\

24-31 #disulfide_bonds\

SUMMARY #length 35 #molecular-weight 3804 #checksum 7419

Query Match 25.08; Score 64; DB 5; Length 35;

Best Local Similarity 45.5%; Pred.No. 1.46e+01;

Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 16 CCRGRPCRCSM 26

QV 6 CCRQKTCSCRL 16

||||:|:|

RESULT 15

ENTRY

TITLE

ORGANISM

REFERENCE

#authors

#submitters

#cross-references PDB:1OMA

COMMENT

COMMENT

FEATURE

10-12,34-36,26-27 #region beta sheet\

13-16 #region turn (type II)\

4-20 #disulfide_bonds\

12-25 #disulfide_bonds\

19-36 #disulfide_bonds\

27-34 #disulfide_bonds\

SUMMARY #length 48 #molecular-weight 5281 #checksum 7904

Query Match 25.08; Score 64; DB 5; Length 48;

Best Local Similarity 45.5%; Pred.No. 1.46e+01;

Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 19 CCRGRPCRCSM 29

QV 6 CCRQKTCSCRL 16

||||:|:|

Search completed: Thu Jul 30 09:20:24 1998

Job time : 12 secs.

MORFAL (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:21:47 1998; MasPar time 1.31 Seconds
147.949 Million cell updates/sec

ular output not generated.

Title: >US-08-938-548A-8
Description: (1-33) from US08938548A.pap
Perfect Score: 256
Sequence: 1 QPLPDCRCQKTCSCRLYELHAGNHAAGILTL 33

Scoring table: PAM 150
Gap 11

Searched: 63816 seqs, 5850866 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5-COMB 2:PCT9_COMB 3:backfiles

Statistics: Mean 21.451; Variance 82.349; scale 0.260

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match Length	ID	Description
1	66	25.8	39	US-08-249-322A-40
2	66	25.8	39	US-08-446-Sequence 20, Applicatio
3	66	25.8	39	US-08-446-Sequence 20, Applicatio
4	66	25.8	39	PCT-US95-0 Sequence 40, Applicatio
5	66	25.8	39	US-08-469-Sequence 40, Applicatio
6	66	25.8	348	PCT-US95-0 Sequence 2, Applicatio
7	64	25.0	1167	US-08-485-Sequence 6, Applicatio
8	64	25.0	1168	US-08-620-Sequence 9, Applicatio
9	62	24.2	1464	US-08-026-Sequence 1, Applicatio
10	61	23.8	42	US-08-487-Sequence 23, Applicatio
11	61	23.8	42	US-08-480-Sequence 23, Applicatio
12	61	23.8	42	US-08-477-Sequence 23, Applicatio
13	61	23.8	42	US-08-137-Sequence 23, Applicatio
14	61	23.8	59	US-08-233-Sequence 51, Applicatio
15	60	23.4	222	5223425-6 Patent No. 5223425
16	60	23.4	334	US-08-279-Sequence 10, Applicatio
17	59	23.0	283	US-08-658-Sequence 2, Applicatio
18	59	23.0	1019	US-08-296-Sequence 4, Applicatio
19	59	23.0	1083	US-08-296-Sequence 2, Applicatio
20	57	22.3	21	US-08-019-Sequence 5, Applicatio
21	57	22.3	48	US-08-379-Sequence 1, Applicatio
22	57	22.3	724	US-08-121-Sequence 62, Applicatio
23	57	22.3	724	PCT-US94-1 Sequence 62, Applicatio

CC FILING DATE: 03-SEP-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/907,138
CC FILING DATE: 30-JUN-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/863,703
CC FILING DATE: 03-APRIL-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: U.K. 91 07566.3
CC FILING DATE: 10-APRIL-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Hanson, Norman D.
CC REGISTRATION NUMBER: 30,946
CC REFERENCE/DOCKET NUMBER: LUD 5250.5
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 688-9200
CC TELEFAX: (212) 838-3884
CC INFORMATION FOR SEQ ID NO: 40:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 39
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC SEQUENCE 39 AA; 3760 MW; 6521 CN;
CC
CC Query Match 25.8%; Score 66; DB 2; Length 39;
CC Best Local Similarity 70.0%; Pred. No. 2.23e+01;
CC Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
CC
CC Db 8 CCR-TTCACR 16
CC III II:II
CC QY 6 CCRQKTCSCR 15
CC
CC RESULT 4
CC ID PCT-US94-05083C-40 STANDARD; PRT; 39 AA.
CC XX
CC AC xxxxxx
CC XX
CC DT
CC XX
CC DE
CC XX
CC Sequence 40, Application PC/TUS9405083C
CC
CC Sequence 40, Application PC/TUS9405083C
CC GENERAL INFORMATION:
CC APPLICANT: Robert Sklar, Mark Marchionni,
CC APPLICANT: David I. Gwynne
CC TITLE OF INVENTION: METHODS FOR ALTERING
CC TITLE OF INVENTION: MUSCLE CONDITION
CC NUMBER OF SEQUENCES: 185
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson
CC STREET: 225 Franklin Street
CC CITY: Boston
CC STATE: Massachusetts
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 5.25 inch, 360
CC MEDIUM TYPE: kb storage
CC COMPUTER: IBM
CC OPERATING SYSTEM: PC-DOS
CC SOFTWARE: Wordperfect
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/05083C
CC FILING DATE: 06-MAY-94
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/209,204
CC FILING DATE: 08-MAR-94
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/059,022
CC FILING DATE: 06-MAY-93
CC ATTORNEY/AGENT INFORMATION:

CC NAME: Clark, Paul T.
CC REGISTRATION NUMBER: 30,162
CC REFERENCE/DOCKET NUMBER: 04585/028W01
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 542-5070
CC TELEFAX: (617) 542-8906
CC TELEX: 200154
CC INFORMATION FOR SEQ ID NO: 40:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 39
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC SEQUENCE 39 AA; 3760 MW; 6521 CN;
CC
CC Query Match 25.8%; Score 66; DB 2; Length 39;
CC Best Local Similarity 70.0%; Pred. No. 2.23e+01;
CC Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
CC
CC Db 8 CCR-TTCACR 16
CC III II:II
CC QY 6 CCRQKTCSCR 15
CC
CC RESULT 5
CC ID US-08-469-569-40 STANDARD; PRT; 39 AA.
CC XX
CC AC xxxxxx
CC XX
CC DT
CC XX
CC DE
CC XX
CC Sequence 40, Application US/08469569
CC
CC Sequence 40, Application US/08469569
CC Patent No. 5606032
CC GENERAL INFORMATION:
CC APPLICANT: Goodearl, Andrew; Stroobant, Paul;
CC APPLICANT: Minghetti, Luisa; Waterfield, Michael; Marchionni, Mark;
CC APPLICANT: Chen, Maio Su; Hiles, Ian
CC TITLE OF INVENTION: Glial Mitogenic Factors, Their
CC TITLE OF INVENTION: Preparation and Use
CC NUMBER OF SEQUENCES: 184
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Felfe & Lynch
CC STREET: 805 Third Avenue
CC CITY: New York City
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10022
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
CC COMPUTER: IBM
CC OPERATING SYSTEM: PC-DOS
CC SOFTWARE: Wordperfect
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/469,569
CC FILING DATE: 06-JUN-1995
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/036,555
CC FILING DATE: 24-MAR-1993
CC APPLICATION NUMBER: 07/965,173
CC FILING DATE: 23-OCT-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/940,389
CC FILING DATE: 03-SEP-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/907,138
CC FILING DATE: 30-JUN-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/863,703
CC FILING DATE: 03-APRIL-1992
CC PRIOR APPLICATION DATA:

CC REFERENCE/DOCKET NUMBER: MA48DD2.C1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (904) 375-8100
CC TELEFAX: (904) 372-5800
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1167 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1167 AA; 131657 MW; 7095219 CN;
SQ
Query Match 25.0%; Score 64; DB 1; Length 1167;
Best Local Similarity 45.5%; Pred. No. 3.29e+01;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Db 873 PNCCKPACQC 883
1:11: 11
4 PDCCRQKTCSC 14
RESULT 8
ID US-08-620-717A-9 STANDARD; PRT; 1168 AA.
XX
AC xxxxxx
XX
DT
DE
XX
XX
Sequence 9, Application US/08620717A
CC
XX
Sequence 9, Application US/08620717A
CC
Patent No. 5670365
CC
GENERAL INFORMATION:
CC APPLICANT: Feitelson, Jerald S.
CC TITLE OF INVENTION: Identification of, and Uses For, Nematocidal
CC TITLE OF INVENTION: Bacillus thuringiensis Genes, Toxins, and Isolates
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Saliwanchik & Saliwanchik
CC STREET: 2421 N.W. 41st Street, Suite A-1
CC CITY: Gainesville
CC STATE: Florida
CC COUNTRY: USA
CC ZIP: 32606
CC
COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/620,717A
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/540,104
CC FILING DATE: 06-OCT-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Saliwanchik, David R.
CC REGISTRATION NUMBER: 31,794
CC REFERENCE/DOCKET NUMBER: MA94.C1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (352) 375-8100
CC TELEFAX: (352) 372-5800
CC INFORMATION FOR SEQ ID NO: 9:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1168 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC ORIGINAL SOURCE:

CC INDIVIDUAL ISOLATE: 167P
SQ SEQUENCE 1168 AA; 131542 MW; 7175894 CN;
Query Match 25.0%; Score 64; DB 1; Length 1168;
Best Local Similarity 45.5%; Pred. No. 3.29e+01;
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Db 873 PNCCKPACQC 883
1:11: 11
4 PDCCRQKTCSC 14
RESULT 9
ID US-08-026-138E-1 STANDARD; PRT; 1464 AA.
XX
AC xxxxxx
XX
DT
DE
XX
XX
Sequence 1, Application US/08026138E
CC
XX
Sequence 1, Application US/08026138E
CC
Patent No. 5502166
CC
GENERAL INFORMATION:
CC APPLICANT: Masayoshi MISHINA
CC TITLE OF INVENTION: NOVEL PROTEINS AND GENES CODING THE SAME
CC NUMBER OF SEQUENCES: 19
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Nishiohata Residence 1-107
CC STREET: 5214, Nishiohata-machi
CC CITY: Niigata-shi
CC STATE: Niigata-ken
CC COUNTRY: JAPAN
CC ZIP: 951
CC
COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: MS-DOS v.5
CC SOFTWARE: Word Perfect 5.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/026,138E
CC FILING DATE: 26-FEB-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: JP 39563/1992
CC FILING DATE: 26-FEB-1992
CC APPLICATION NUMBER: JP 173155/1992
CC FILING DATE: 30-JUN-1992
CC APPLICATION NUMBER: JP 215017/1992
CC FILING DATE: 12-AUG-1992
CC APPLICATION NUMBER: JP 303878/1992
CC FILING DATE: 13-NOV-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Hamburg, C. Bruce
CC REGISTRATION NUMBER: 22,389
CC REFERENCE/DOCKET NUMBER: F-4551
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 986-2340
CC TELEFAX: (212) 953-7733
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1464 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single strand
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC ORIGINAL SOURCE:
CC ORGANISM: mouse
CC TISSUE TYPE: cerebellum
CC PUBLICATION INFORMATION:
CC AUTHORS: Masayoshi MISHINA
CC TITLE: NOVEL PROTEINS AND GENES CODING THE SAME
CC RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 to 1464
CC SEQUENCE 1464 AA; 165489 MW; 11224000 CN;
SQ

```

RESULT 12
ID US-08-477-383-23 STANDARD; PRT; 42 AA.
XX
AC
XX
DT
XX
DE
XX
Sequence 23, Application US/08477383
Sequence 23, Application US/08477383
Patent No. 5589340
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hilliard, David R.
APPLICANT: Macintosh, J. Michael
APPLICANT: Santos, Ameurfin D.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue, N.W., Suite 1000
CITY: Washington
STATE: DC
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,383
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/137,800
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/084,848
FILING DATE: 29-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-107673
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Conus ochroleucus
SEQUENCE 42 AA; 4618 MW; 9505 CN;

Query Match 23.8%; Score 61; DB 1; Length 42;
Best Local Similarity 47.4%; Pred. No. 5.85e+01;
Matches 9; Conservative 3; Mismatches 5; Indels 2; Gaps 2;

Db 14 PVPSCCQVSSCW-NLYGL 31
|:| |:| |:| |:|
QY 2 PLP-DCCROKTCSCRLYL 19

RESULT 13
ID US-08-137-800-23 STANDARD; PRT; 42 AA.
XX
AC
XX
DT
XX
DE
XX
Sequence 23, Application US/08137800
Sequence 23, Application US/08137800
Patent No. 5514774
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hilliard, David R.
APPLICANT: Macintosh, J. Michael
APPLICANT: Santos, Ameurfin D.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue N.W., Suite 1000
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137,800
FILING DATE: 19-OCT-1993
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-104763
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Conus ochroleucus
SEQUENCE 42 AA; 4618 MW; 9505 CN;

Query Match 23.8%; Score 61; DB 1; Length 42;
Best Local Similarity 47.4%; Pred. No. 5.85e+01;
Matches 9; Conservative 3; Mismatches 5; Indels 2; Gaps 2;

Db 14 PVPSCCQVSSCW-NLYGL 31
|:| |:| |:| |:|
QY 2 PLP-DCCROKTCSCRLYL 19

Sequence 51, Application US/08233788A
Sequence 51, Application US/08233788A
Patent No. 5635617
GENERAL INFORMATION:
APPLICANT: Doran, James L.

```

```

XX
AC
XX
DT
XX
DE
XX
Sequence 23, Application US/08137800
Sequence 23, Application US/08137800
Patent No. 5514774
GENERAL INFORMATION:
APPLICANT: Olivera, Baldomero M.
APPLICANT: Cruz, Lourdes J.
APPLICANT: Hilliard, David R.
APPLICANT: Macintosh, J. Michael
APPLICANT: Santos, Ameurfin D.
TITLE OF INVENTION: Conotoxin Peptides
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue N.W., Suite 1000
CITY: Washington
STATE: DC
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/137,800
FILING DATE: 19-OCT-1993
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ihnen, Jeffrey L.
REGISTRATION NUMBER: 28,957
REFERENCE/DOCKET NUMBER: 24260-104763
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4810
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Conus ochroleucus
SEQUENCE 42 AA; 4618 MW; 9505 CN;

Query Match 23.8%; Score 61; DB 1; Length 42;
Best Local Similarity 47.4%; Pred. No. 5.85e+01;
Matches 9; Conservative 3; Mismatches 5; Indels 2; Gaps 2;

Db 14 PVPSCCQVSSCW-NLYGL 31
|:| |:| |:| |:|
QY 2 PLP-DCCROKTCSCRLYL 19

RESULT 14
ID US-08-233-788A-51 STANDARD; PRT; 59 AA.
XX
AC
XX
DT
XX
DE
XX
Sequence 51, Application US/08233788A
Sequence 51, Application US/08233788A
Patent No. 5635617
GENERAL INFORMATION:
APPLICANT: Doran, James L.

```

WQREH (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Thu Jul 30 09:19:34 1998; Maspar time 3.95 Seconds
128.266 Million cell updates/sec
ular output not generated.

Title: >US-08-938-548A-8
Description: (1-33) from US08938548A.pap
Perfect Score: 256
Sequence: 1 QPLPDCRQKTCSCRLYELLHGAGNHAAGILTL 33

Scoring table: PAM 150
Gap 11

Searched: 124785 seqs, 15338987 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq31-2
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27

Statistics: Mean 22.991; Variance 84.736; scale 0.271

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	77	30.1	363	23	W12414	Porcine complement in
2	75	29.3	373	18	R98461	Murine ICE-ced-3 homo
3	75	29.3	373	13	R66767	Murine interleukin-1
4	69	27.0	329	26	W29877	Lysophosphatidic acid
5	66	25.8	102	1	P95679	Xenopus Vg1 protein f
6	66	25.8	348	20	W01619	Human uridine diphosph
7	65	25.4	78	5	P20020	Sequence of a foot an
8	64	25.0	45	23	W10106	Human agouti signalli
9	64	25.0	45	23	W10105	Murine agouti signall
10	64	25.0	48	9	R45611	AG1 toxin.
11	64	25.0	48	11	R60293	Calcium channel inhib
12	64	25.0	48	8	R44209	A. aperta venom fract
13	64	25.0	130	23	W10102	Human agouti signalli
14	64	25.0	131	23	W10101	Murine agouti signall
15	64	25.0	1167	27	W31504	Nematode toxin 167 p
16	64	25.0	1167	20	W10653	Bacillus thuringiensis
17	64	25.0	1168	23	W16326	Nematocidal toxin 167
18	62	24.2	34	10	R55088	Tarantula spider veno
19	62	24.2	1464	10	R55529	Human NMDA R2A recept

20	62	24.2	1464	12	R66039	Human N-methyl-D-aspa
21	62	24.2	1464	18	R80970	Human excitatory amin
22	62	24.2	1464	8	R42054	Glutamic acid recepto
23	62	24.2	1464	13	R44192	Rat NMDA receptor sub
24	61	23.8	42	20	W12745	A-lineage conotoxin p
25	61	23.8	42	23	W24890	Predatory cone snail
26	61	23.8	59	23	W23579	Salmonella enteritidi
27	61	23.8	253	24	W27333	Human membrane antige
28	61	23.8	374	12	R62758	TcTA sequence.
29	61	23.8	3672	27	W31950	Human bg protein asso
30	61	23.8	3801	27	W31949	Human bg protein asso
31	60	23.4	140	5	R27559	Cyn d allergen B1.
32	60	23.4	147	26	W23671	Cherry polyphenol oxi
33	60	23.4	334	26	W36071	E. coli DNA polymeras
34	60	23.4	334	8	R40126	DNA polymerase III ho
35	60	23.4	367	8	R40115	APP-HCV-E2 fusion pro
36	59	23.0	39	1	R05633	Spider venom peptide
37	59	23.0	44	2	R11517	N-terminal sequence o
38	59	23.0	48	8	R42948	Funnel-web spider ven
39	59	23.0	283	27	W37469	Connexin-32.
40	58	22.7	49	3	P91100	Sequence of viper ven
41	58	22.7	49	3	P91095	Sequence of viper ven
42	58	22.7	129	2	R10038	Ovine FSH beta subuni
43	58	22.7	334	16	R92020	Ikaros protein.
44	58	22.7	2186	27	W31948	Mouse 22B/30B (candid
45	58	22.7	3788	25	W23594	Murine Lystl long iso

ALIGNMENTS

RESULT 1
ID W12414 standard; Protein; 363 AA.
AC W12414;
DT 24-SEP-1997 (first entry)
DE Porcine complement inhibitor.
KW porcine; pig; complement; inhibitor; organ transplantation;
KW analysis; promoter.
OS Sus scrofa.
PN W09700951-AL.
PD 09-JAN-1997.
PF 19-JUN-1996; JP-178254.
PR 20-JUN-1995; JP-178254.
PA (NIME-) NIPPON MEAT PACKERS INC.
PA (NIHA-) NIPPON HAM KK.
PI Murakami H, Shigehisa T, Toyomura K;
DR WPI; 97-087378/08.
DR N-PSDB; T61098.
PT DNA encoding porcine complement inhibitor - useful in porcine organ
transplant to humans
PS Claim 3; Page 12-14; 20pp; Japanese.
CC This protein is a porcine complement inhibitor encoded by pmcpDNA
(T61098). The DNA is useful for large scale production of
CC recombinant porcine complement inhibitor, which is useful for
CC porcine organ transplantation into humans. The DNA clone pmcpDNA is
CC also useful in the analysis of the promoter region of porcine complement
inhibitor.
SQ Sequence 363 AA;
Query Match 30.1%; Score 77; DB 23; Length 363;
Best Local Similarity 39.3%; Pred. No. 6.26e+00;
Matches 11; Conservative 7; Mismatches 9; Indels 1; Gaps 1;
Db 98 plqacrrkacs-nlpdplngqvssyng 124
|||:|||||:|:|:|:
QY 2 PLPDCRQKTCSCRLYELLHGAGNHAAG 29

RESULT 2
ID R98461 standard; Protein; 373 AA.
AC R98461;
DT 25-SEP-1996 (first entry)
DE Murine ICE-ced-3 homologue.
KW mich-2; murine ICE-ced-3 homologue; programmed cell death;

CC congenital craniofacial and other skeletal or dental anomalies, induction
 CC of local endochondral bone formation in non-union fractures, peridental
 CC applns. requiring bone formation and cartilage repair, eg in the
 CC treatment of osteoarthritis.
 CC See also P95679-P95692 and N95097.
 SQ Sequence 102 AA;

Query Match 25.8%; Score 66; DB 1; Length 102;
 Best Local Similarity 56.3%; Pred. No. 5.80e+01;

Matches 9; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

Db 34 cpyplieilngs-nha 48

QY 12 CSCRYELLHGAGNHA 27

RESULT 6

ID W01619 standard; Protein; 348 AA.

AC W01619;

DE 24-APR-1997 (first entry)

Human uridine diphosphate galactose-4-epimerase.

Uridine diphosphate galactose-4-epimerase; UDP-G4E; galactosaemia;

therapy; diagnosis.

OS Homo sapiens.

PN W09635778-A1.

PD 14-NOV-1996.

PF 11-MAY-1995; U05785.

PR 11-MAY-1995; WO-005785.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ji H, Rosen CA;

WPI: 96-518666/51.

DR N-PSDB; T59301.

PT DNA encoding human uridine diphosphate galactose-4-epimerase - used

in the treatment and diagnosis of galactosaemia

PS Claim 1; Page 43-44; 59pp; English.

CC Human mature uridine diphosphate galactose-4-epimerase (UDP-G4E)

UDP-galactose that allows galactose residues to enter into the main

pathways of glucose metabolism. A deficiency of the enzyme results

in galactosaemia. The amino acid sequence of UDP-G4E was deduced

from a cDNA clone (T58301) derived from a human endometrial tumour

library. Recombinant UDP-G4E polypeptides can be produced in

transformed host (e.g. E. coli, COS, Sf9 insect) cells or expressed

in vivo for use in the treatment of UDP-G4E deficiency, e.g.

galactosaemia.

SQ Sequence 348 AA;

Query Match 25.8%; Score 66; DB 20; Length 348;
 Best Local Similarity 53.8%; Pred. No. 5.80e+01;

Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Db 262 cgcrynlgtgtg 274

QY 12 CSCRYELLHGAG 24

RESULT 7

ID P20020 standard; Protein; 78 AA.

AC P20020;

DE 20-AUG-1992 (first entry)

Sequence of a foot and mouth disease virus capsid protein

DE encoded by a region of recombinant plasmid pFA61/t76

KW Vaccine; antibody; capsid protein; immunogen; antigen;

foot and mouth disease.

OS Foot and mouth disease virus.

PN EP--48455-A.

PD 31-MAR-1982.

PF 17-SEP-1981.

PR 18-SEP-1980; GB-030208.

PR 22-OCT-1980; GB-034130.

PR 27-NOV-1980; GB-038147.

PR 08-APR-1981; GB-011064.

PR 18-AUG-1981; GB-025150.

PA (NATR) National Res Dev Corp.
 PA (WELL) Wellcome Foundation Ltd.
 PI Boothroyd JC, Cross GAM, Highfield PE, Winther MD, Rowlands DJ,
 PI Brown F, Harris TUR, Lowe PA;
 DR WPI: 82-26702E/14.
 DR N-PSDB; N20019.
 PT DNA corresp. to (part of) foot and mouth disease virus RNA - useful
 in prepn. of vaccines for producing antibodies against the virus
 PS Example; Fig 6; 57pp; English.
 CC The inventors claim a DNA molecule comprising a nucleotide sequence
 corresp. to all or a portion of foot-and-mouth disease virus RNA
 (FMDV). The DNA molecule is esp. for a precursor of FMDV capsid
 protein. It esp. codes for FMDV protein p88 and VP1-VP4. It may code
 for VP4, VP2, VP3 and VP1 contiguously. The inventors also claim a
 vaccine for stimulating prodn. of antibodies against FMDV in a
 mammal which comprises at least one of the above recombinant
 CC proteins produced by a host cell transformed with the DNA.
 SQ Sequence 78 AA;

Query Match 25.4%; Score 65; DB 5; Length 78;
 Best Local Similarity 37.5%; Pred. No. 7.06e+01;

Matches 9; Conservative 9; Mismatches 5; Indels 1; Gaps 1;

Db 45 qntcsthtygglbst-thstlvls 67

QY 9 QKTCSCRYELLHGAGNHAAGILT 32

RESULT 8

ID W10106 standard; protein; 45 AA.

AC W10106;

DE 18-SEP-1997 (first entry)

Human agouti signalling protein fragment #2.

KW Agouti signalling protein; ASP; depigmenting activity; cosmetic;

KW hyperpigmentation condition; melasma photoageing spots; solar keratosis;

KW post-inflammatory hyperpigmentation; wound healing; eumelanogenesis;

KW vitiligo; leucoderma; albinism; hair greying.

OS Homo sapiens.

PN W09700892-A2.

PD 09-JAN-1997.

PF 21-JUN-1996; U10695.

PR 23-JUN-1995; US-000436.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Hearing VJ;

DR WPI: 97-087323/08.

PT New agouti signal protein peptide(s) and nucleic acids - used for

altering melanin prodn., for treating e.g. melasma photo-ageing

PT spots, solar keratosis or vitiligo

PS Claim 10; Page 11; 67pp; English.

CC The sequences given in W10101-29 are biologically active peptides

and fragments of the agouti signalling protein (ASP) which have

depigmenting activity. These peptides are useful for cosmetic purposes

and for clinical application in the prevention or treatment of various

hyperpigmentary conditions and diseases such as melasma photoageing

spots, solar keratosis, and post-inflammatory hyperpigmentation such as

occurs at sites of wound healing. They can also be used to provide

enhanced eumelanogenesis for treating eg. vitiligo, leucoderma, some

forms of albinism and hair greying.

SQ Sequence 45 AA;

Query Match 25.0%; Score 64; DB 23; Length 45;
 Best Local Similarity 41.2%; Pred. No. 8.59e+01;

Matches 7; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

Db 16 papaccdpccascqcrff 32

QY 2 PLPDCCRQ-KTCSRLY 17

RESULT 9

ID W10105 standard; protein; 45 AA.

AC W10105;

DT 17-SEP-1997 (first entry)

PS Example; Page 17; 28pp; English.
 CC The sequence is that of a polypeptide present in fraction K of the
 CC venom of *Agelenopsis aperta*, it blocks calcium channels in cells
 CC of both mammals and invertebrates, partic. those affecting neuronal
 CC and muscle cells. It may be used in the treatment of angina,
 CC hypertension, cardiomyopathies, supraventricular arrhythmia,
 CC oesophageal achalasia, premature labour, and Raynaud's disease.
 CC It may also be of use in the study of cell physiology and in the
 CC control of invertebrate pests. It may be produced synthetically.
 SQ Sequence 48 AA;

Query Match 25.0%; Score 64; DB 8; Length 48;

Best Local Similarity 45.5%; Pred. No. 8.59e+01;

Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

DB 19 ccrgprcscms 29

QY 6 CCRQKTCSCRL 16

BLT 13

W10102 standard; protein; 130 AA.

AC W10102;

DE Human agouti signalling protein.

KW Agouti signalling protein; ASP; depigmenting activity; cosmetic;

KW hyperpigmentary condition; melasma photoageing spots; solar keratosis;

KW post-inflammatory hyperpigmentation; wound healing; eumelanogenesis;

KW vitiligo; leucoderma; albinism; hair greying.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Peptide 1..21

FT region /note= "Signal sequence"

FT region /note= "Start of basic region"

FT region /note= "Start of Cysteine-rich motif"

PN W09700892-A2.

PD 09-JAN-1997.

PF 21-JUN-1996; U10695.

PR 23-JUN-1995; US-000436.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Hearing VJ;

DR WPI: 97-087323/08.

PT New agouti signal protein peptide(s) and nucleic acids - used for

PT altering melanin prodn., for treating e.g. melasma photo-ageing

PT spots, solar keratosis or vitiligo

PS Claim 6; Page 8-9; 67pp; English.

CC The sequences given in W10101-29 are biologically active peptides

CC and fragments of the agouti signalling protein (ASP) which have

CC depigmenting activity. These peptides are useful for cosmetic purposes

CC and for clinical application in the prevention or treatment of various

CC hyperpigmentary conditions and diseases such as melasma photoageing

CC spots, solar keratosis, and post-inflammatory hyperpigmentation such as

CC occurs at sites of wound healing. They can also be used to provide

CC enhanced eumelanogenesis for treating eg. vitiligo, leucoderma, some

CC forms of albinism and hair greying.

CC Sequence 130 AA;

Query Match 25.0%; Score 64; DB 23; Length 130;

Best Local Similarity 41.2%; Pred. No. 8.59e+01;

Matches 7; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

DB 101 papaccdpascqrcff 117

QY 2 PLPDCCRQ-KTCSCLY 17

RESULT 14

ID W10101 standard; protein; 131 AA.

AC W10101;

DE Murine agouti signalling protein.

KW Agouti signalling protein; ASP; depigmenting activity; cosmetic;
 KW hyperpigmentary condition; melasma photoageing spots; solar keratosis;
 KW post-inflammatory hyperpigmentation; wound healing; eumelanogenesis;
 KW vitiligo; leucoderma; albinism; hair greying.
 OS Mus musculus.

PH Key Location/Qualifiers

FT Peptide 1..21

FT region /note= "Signal sequence"

FT region /note= "Start of basic region"

FT region /note= "Start of Cysteine-rich motif"

PN W09700892-A2.

PD 09-JAN-1997.

PF 21-JUN-1996; U10695.

PR 23-JUN-1995; US-000436.

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

PI Hearing VJ;

DR WPI: 97-087323/08.

PT New agouti signal protein peptide(s) and nucleic acids - used for

PT altering melanin prodn., for treating e.g. melasma photo-ageing

PT spots, solar keratosis or vitiligo

PS Claim 5; Page 8-9; 67pp; English.

CC The sequences given in W10101-29 are biologically active peptides

CC and fragments of the agouti signalling protein (ASP) which have

CC depigmenting activity. These peptides are useful for cosmetic purposes

CC and for clinical application in the prevention or treatment of various

CC hyperpigmentary conditions and diseases such as melasma photoageing

CC spots, solar keratosis, and post-inflammatory hyperpigmentation such as

CC occurs at sites of wound healing. They can also be used to provide

CC enhanced eumelanogenesis for treating eg. vitiligo, leucoderma, some

CC forms of albinism and hair greying.

CC Sequence 131 AA;

Query Match 25.0%; Score 64; DB 23; Length 131;

Best Local Similarity 41.2%; Pred. No. 8.59e+01;

Matches 7; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

DB 102 papaccdpascqrcff 118

QY 2 PLPDCCRQ-KTCSCLY 17

RESULT 15

ID W31504 standard; Protein; 1167 AA.

AC W31504;

DE Nematoxin (first entry)

KW PCR primer; amplify; nematode toxic protein; *Bacillus thuringiensis*;

KW delta-endotoxin gene; nematode pest control; *Panagrellus redivivus*;

KW 167P protein.

OS *Bacillus thuringiensis*.

PN W09734926-A2.

PD 25-SEP-1997.

PF 21-MAR-1997; U04755.

PR 21-MAR-1996; US-590554.

PA (MYCO) MYCOGEN CORP.

PI Fu J, Narva KE, Payne J;

DR WPI: 97-480163/44.

DR N-PSDB: T89185.

PT *Bacillus thuringiensis* toxin gene - useful in recombinant hosts,

PT particularly plants for the control of nematodes

PS Claim 4; Page 35-39; 44pp; English.

CC This sequence represents the protein encoded by a polynucleotide of the

CC invention. The polynucleotide of the invention is a sequence from a

CC *Bacillus thuringiensis* (Bt) isolate selected from PS80J11, PS158D5, a

CC PS167P, PS169E, PS177F1, PS177G, PS204G4 and PS204G6, that encodes a

CC toxin active against nematodes. This sequence represents the 167P

CC protein, and is a delta-endotoxin protein. The polynucleotides and toxins

CC can be used for the control of nematode pests such as *Panagrellus*

CC *redivivus*. 1167 AA;

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm
on: Thu Jul 30 09:17:23 1998; MasPar time 9.36 Seconds
Molecular output not generated.
Title: >US-08-938-548A-6
Description: (1-130) from US08938548A.pep
Perfect Score: 954
Sequence: 1 MNPSTKVPWNAVTLILL.....GRRCPATATAPRGGSRV 130
Scoring table: PAM 150
Gap 11
Searched: 140542 seqs, 42109429 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: spiremb15
1:sp_fungi 2:sp_human 3:sp_invertebrate 4:sp_mammal
5:sp_mhc 6:sp_organelle 7:sp_phase 8:sp_plant
9:sp_bacteria 10:sp_rodent 11:sp_virus 12:sp_vertebrate
13:sp_unclassified
Statistics: Mean 39.989; Variance 88.114; scale 0.454
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	106	11.1	679	8	HYPOTHETICAL 74.6 KD P	3.56e-02
2	103	10.8	567	10	PLATELET GLYCOPROTEIN	8.61e-02
3	98	10.3	574	10	ACTOXYACIL HYDROLASE	3.65e-01
4	98	10.3	641	10	FRIZZLED PROTEIN HOMOL	3.65e-01
5	92	9.6	106	8	PROLINE- AND LEUCINE-R	1.95e+00
6	91	9.5	440	10	LECITHIN:CHOLESTEROL A	2.56e+00
7	91	9.5	602	10	CYCLOOXYGENASE 1	2.56e+00
8	91	9.5	602	10	CYCLOOXYGENASE ISOFORM	2.56e+00
9	90	9.4	833	2	SORTILIN PRECURSOR	3.36e+00
10	89	9.3	154	2	COL1A1 AND PDGFB FUSIO	4.40e+00
11	89	9.3	185	2	C-SIS PROTO-ONCOGENE (4.40e+00
12	89	9.3	331	12	LEUCINE ZIPPER WITH BA	4.40e+00
13	89	9.3	602	10	PROSTAGLANDIN H SYNTHA	4.40e+00
14	88	9.2	226	9	HYPOTHETICAL 23.1 KD P	5.75e+00
15	88	9.2	1174	4	TIGHT JUNCTION PROTEIN	5.75e+00
16	87	9.1	225	11	POLYPROTEIN PRECURSOR	7.50e+00
17	87	9.1	271	11	DNA BINDING PROTEIN E1	7.50e+00
18	87	9.1	593	12	PROSTAGLANDIN ENDOPERO	7.50e+00
19	87	9.1	599	2	Q15122	7.50e+00
20	87	9.1	1238	3	SIMILARITY TO MOUSE SM	7.50e+00

21	86	9.0	1099	10	P97527	NB-2.	9.76e+00
22	86	9.0	1358	11	Q65813	(STRAIN OREGON) P125 (9.76e+00
23	86	9.0	1839	9	Q30765	POLYKETIDE SYNTHASE MO	9.76e+00
24	86	9.0	4340	9	Q30764	POLYKETIDE SYNTHASE MO	9.76e+00
25	85	8.9	341	2	O15354	G PROTEIN COUPLED RECE	1.27e+01
26	85	8.9	363	4	O02839	PORCINE MEMBRANE COFAC	1.27e+01
27	85	8.9	613	2	O00348	PUTATIVE G PROTEIN-COU	1.27e+01
28	85	8.9	613	2	O14768	PUTATIVE G PROTEIN-COU	1.27e+01
29	85	8.9	878	11	Q83101	RNA-DEPENDENT RNA POLY	1.27e+01
30	84	8.8	84	10	Q60471	ANION EXCHANGER ISOFOR	1.64e+01
31	84	8.8	102	8	Q43495	108 PROTEIN PRECURSOR	1.64e+01
32	84	8.8	335	12	Q91654	THYROID HORMONE INDUCE	1.64e+01
33	84	8.8	342	6	Q35413	NADH DEHYDROGENASE SUB	1.64e+01
34	84	8.8	701	11	Q65568	HYPOTHETICAL 72.6 KD P	1.64e+01
35	84	8.8	729	10	Q60470	ANION EXCHANGER 2 A (F	1.64e+01
36	84	8.8	1117	2	O00542	DIACYLGLYCEROL KINASE	1.64e+01
37	83	8.7	127	2	Q15016	ORF, COMPLETE CDS (FRA	2.12e+01
38	83	8.7	320	3	Q17956	M01B2.2.	2.12e+01
39	83	8.7	329	3	Q27916	MYOMODULIN (FRAGMENT).	2.12e+01
40	83	8.7	356	9	P76081	FROM BASES 1449468 TO	2.12e+01
41	83	8.7	370	3	Q07974	MYOMODULIN PRECURSOR (2.12e+01
42	83	8.7	426	9	P77233	FERRDOXIN.	2.12e+01
43	83	8.7	498	3	Q17021	MALTASE-LIKE PROTEIN A	2.12e+01
44	83	8.7	551	9	P72405	PCBR.	2.12e+01
45	83	8.7	1109	4	O19179	GUANYLATE CYCLASE E.	2.12e+01

ALIGNMENTS

RESULT 1	PRELIMINARY; PRT; 679 AA.
ID O23352	
AC O23352	
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)	
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)	
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)	
DE HYPOTHETICAL 74.6 KD PROTEIN	
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).	
OC EUKARYOTA; PLANTA; EMERYOPHYTA; ANGIOSPERMAE; DICOITYLEDONEAE;	
OC CAPPARALES; CRUCIFERAE.	
RP SEQUENCE FROM N.A.	
RA BEVAN M., STIEKEMA W., MURPHY G., WAMBUIT R., POHL T., TERRY N.,	
RA KREIS M., KAVANAGH T., ENTIAN K.D., RIEGER M., JAMES R.,	
RA PUIGDOMENECH P., HATZOPOULOS P., OBERWAITER B., DUESTERHOFT A., JONES J.,	
RA PALME K., ANSGORGE W., DELSENY M., BANCROFT I., MEWES H.W., SCHUELLER C.,	
RA CHALWATZIS N.	
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.	
RN [2]	
RP SEQUENCE FROM N.A.	
RA EU ARABIDOPSIS SEQUENCING PROJECT, ESSA;	
RL SUBMITTED (JUN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.	
DR EMBL; Z97337; E326841; -	
KW HYPOTHETICAL PROTEIN.	
SQ SEQUENCE 679 AA; 74635 MW; B301B713 CRC32;	
Query Match 11.1%; Score 106; DB 8; Length 679;	
Best Local Similarity 32.3%; Pred. No. 3.56e-02;	
Matches 20; Conservative 16; Mismatches 22; Indels 4; Gaps 4;	
Db 598 KYCRSK-YETIHQNDNADVLELAIKREMPAELL-R-ASLRHNEEDQNFNLNVRSA 654	
QY 42 KTCSCRLYELLHGAG-NHAAGILTGLKRRPPGLQGLRQLRQLQANGNHAAGILTMGRR 100	
Db 655 SP 656	
QY 101 GA 102	
RESULT 2	
ID O08770	
AC O08770	
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)	
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)	

OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY AND WISTAR; TISSUE=LIVER;
RX MEDLINE; 97363611.
RA WANG J., GEBRE A.K., ANDERSON R.A., PARKS J.S.;
RL BIOCHIM. BIOPHYS. ACTA 1346:207-211(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY AND WISTAR; TISSUE=LIVER;
RA WANG J., ANDERSON R.A., PARKS J.S.;
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + STEROL = STEROL ESTER +
CC
CC
DR EMBL; U62803; G2306762;
DR PROSITE; PS00120; LIPASE_SER; 1.
KW TRANSFERASE; ACYLTRANSFERASE.
SQ SEQUENCE 440 AA; 49882 MW; 6CC16087 CRC32;
Query Match 9.5%; Score 91; DB 10; Length 440;
Best Local Similarity 62.5%; Pred. No. 2.56e+00;
Matches 15; Conservative 2; Mismatches 5; Indels 2; Gaps 1;
Db 1 MCLPGS--PQWVLLLLGLLLPPA 22
QY 1 MNLPTKVPWAATVLLLLLLPPA 24
RESULT 7
ID Q63684
AC Q63684; PRELIMINARY; PRT; 602 AA.
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE CYCLOXYGENASE 1.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-FISHER 344;
RA FENG L., SUN W., XIA Y., TANG W.W., CHANNUGAM P., SOYCOOLA E.,
RA WILSON C.B., HWANG D.;
RL ARCH. BIOCHEM. BIOPHYS. 307:0-0(0).
DR EMBL; U03388; G415638;
SQ SEQUENCE 602 AA; 69115 MW; DC2236E9 CRC32;
Query Match 9.5%; Score 91; DB 10; Length 602;
Best Local Similarity 40.0%; Pred. No. 2.56e+00;
Matches 16; Conservative 7; Mismatches 15; Indels 2; Gaps 2;
Db 1 MSRRSLQFPLLLLLPPPPVLLTDAGVPSVIP-CC 39
QY 1 MNLPTKVPWAATVLLLLLLPPALLS-LGVDAQPLPDC 39
RESULT 8
ID Q63921
AC Q63921; PRELIMINARY; PRT; 602 AA.
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DE CYCLOXYGENASE ISOFORM COX-1 (FRAGMENT).
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94099619.
RA FENG L., SUN W., XIA Y., TANG W.W., CHANNUGAM P., SOYCOOLA E.,
RA WILSON C.B., HWANG D.;

RL ARCH. BIOCHEM. BIOPHYS. 307:361-368(1993).
DR EMBL; S67721; G460556;
FT NON_TER 602
SQ SEQUENCE 602 AA; 69158 MW; 4EBDC921 CRC32;
Query Match 9.5%; Score 91; DB 10; Length 602;
Best Local Similarity 40.0%; Pred. No. 2.56e+00;
Matches 16; Conservative 7; Mismatches 15; Indels 2; Gaps 2;
Db 1 MSRRSLQFPLLLLLPPPPVLLTDAGVPSVIP-CC 39
QY 1 MNLPTKVPWAATVLLLLLLPPALLS-LGVDAQPLPDC 39
RESULT 9
ID Q99523
AC Q99523; PRELIMINARY; PRT; 833 AA.
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE SORTILIN PRECURSOR.
GN SORTIL.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA PETERSEN C.M., NIELSEN M.S., JACOBSEN L., TOMMERUP N., HOLM H.,
RA ROIGAARD H., GLIEMANN J., MADSEN P., MOESTRUP S.K.;
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; X98248; E246784;
KW SIGNAL.
FT SIGNAL. 1 33 POTENTIAL.
FT CHAIN 34 833 SORTILIN.
SQ SEQUENCE 833 AA; 92408 MW; D2E351B9 CRC32;
Query Match 9.4%; Score 90; DB 2; Length 833;
Best Local Similarity 53.3%; Pred. No. 3.36e+00;
Matches 16; Conservative 5; Mismatches 6; Indels 3; Gaps 3;
Db 14 WPHGLGLLLQLPPSTLSQDRLDAPPPP 43
QY 10 WA-ATVLLLL-LLPALLSLG-VDAQPLP 36
RESULT 10
ID O15186
AC O15186; PRELIMINARY; PRT; 154 AA.
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE COLI1 AND POCFB FUSION TRANSCRIPT (FRAGMENT).
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA O'BRIEN K.P.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; Y08643; E293780;
FT NON_TER 1 154
SQ SEQUENCE 154 AA; 16512 MW; 6DD0301B CRC32;
Query Match 9.3%; Score 89; DB 2; Length 154;
Best Local Similarity 43.6%; Pred. No. 4.40e+00;
Matches 17; Conservative 11; Mismatches 6; Indels 5; Gaps 5;
Db 69 LORLLHGDPEEDGAELDINMTRSHSGGEISLA-RGRR 106
QY 79 LORLLQAN-GNH-AAGI-LTMGR-RAGAELEPPCPGRR 113

Search completed: Thu Jul 30 09:18:04 1998
Job time : 41 secs.

MPERCH_PP

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPERCH_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:16:51 1998; MasPar time 5.05 Seconds
646.000 Million cell updates/sec
Full output not generated.

Title: >US-08-938-548A-6
Description: (1-130) from US08938548A.pep
Perfect Score: 954
Sequence: 1 MNLPSYKVPWNAVTL LLLLL.....GRRCPATATAPRGGSRV 130

Scoring table: PAM 150
Gap 11

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 41.891; Variance 75.823; scale 0.552

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	101	10.6	131	1	SECR_PIG	1.78e-02
2	96	10.1	438	1	LCAT_MOUSE	9.43e-02
3	96	10.1	566	1	FBLA_HUMAN	9.43e-02
4	96	10.1	601	1	FBLB_HUMAN	9.43e-02
5	96	10.1	683	1	FBLC_HUMAN	9.43e-02
6	96	10.1	703	1	FBLD_HUMAN	9.43e-02
7	93	9.7	1027	1	CAFF_RIFPA	2.50e-01
8	93	9.7	1061	1	ANPA_HUMAN	2.50e-01
9	91	9.5	251	1	C10B_HUMAN	4.74e-01
10	91	9.5	440	1	LCAT_RAT	4.74e-01
11	91	9.5	440	1	LCAT_HUMAN	4.74e-01
12	89	9.3	241	1	PDGB_HUMAN	8.89e-01
13	89	9.3	331	1	PRP1_HUMAN	1.21e+00
14	88	9.2	252	1	ICP3_HSV1D	1.21e+00
15	88	9.2	254	1	4-1BB_HUMAN	1.21e+00
16	88	9.2	497	1	SC14_YARLI	1.21e+00
17	87	9.1	226	1	TSIS_SWSA	1.65e+00
18	86	9.0	245	1	ICP3_HSV1N	2.24e+00
19	86	9.0	322	1	YCEC_HAEIN	2.24e+00
20	86	9.0	334	1	FEPD_ECOLI	2.24e+00
21	86	9.0	440	1	LCAT_ECAN	2.24e+00
22	86	9.0	1103	1	CYGD_HUMAN	2.24e+00
23	86	9.0	1663	1	CO3_RAT	2.24e+00

24	85	8.9	236	1	PLCL_BOVIN	PLACENTAL LACTOGEN I P	3.03e+00
25	85	8.9	241	1	PDGB_MOUSE	PLATELET-DERIVED GROWT	3.03e+00
26	85	8.9	255	1	YPEL_RHOU	HYPOTHETICAL 28 KD PRO	3.03e+00
27	85	8.9	317	1	LIP1_PSYIM	LIPASE 1 PRECURSOR (EC	3.03e+00
28	84	8.8	90	1	VGE_BPPHX	LYSIS PROTEIN.	4.09e+00
29	84	8.8	90	1	VGE_BPS13	E PROTEIN.	4.09e+00
30	84	8.8	238	1	EFA3_HUMAN	EPHRIN-A3 PRECURSOR (E	4.09e+00
31	84	8.8	315	1	LIP3_MORSP	LIPASE 3 PRECURSOR (EC	4.09e+00
32	84	8.8	319	1	YCEC_ECOLI	HYPOTHETICAL 36.0 KD P	4.09e+00
33	84	8.8	384	1	RN_DROME	GNPASE ACTIVATING PROT	4.09e+00
34	84	8.8	696	1	LSHR_PIG	LUTROPIN-CHORIOGONADOT	4.09e+00
35	84	8.8	1001	1	PTPX_MOUSE	PROTEIN-TYROSINE PHOSP	4.09e+00
36	83	8.7	208	1	GPBB_PAPCY	PLATELET GLYCOPROTEIN	5.50e+00
37	83	8.7	224	1	OXO2_HORVU	OXALATE OXIDASE PRECUR	5.50e+00
38	83	8.7	235	1	FL3L_HUMAN	SL CYTOKINE PRECURSOR	5.50e+00
39	83	8.7	281	1	POTI_ECOLI	PUTRESCINE TRANSPORT S	5.50e+00
40	83	8.7	440	1	LCAT_RABIT	PHOSPHATIDYLCHOLINE-ST	5.50e+00
41	83	8.7	536	1	YABK_ECOLI	HYPOTHETICAL 59.6 KD P	5.50e+00
42	83	8.7	615	1	ALBU_CHICK	SERUM ALBUMIN PRECURSO	5.50e+00
43	83	8.7	1004	1	PTPX_RAT	PROTEIN-TYROSINE PHOSP	5.50e+00
44	83	8.7	1310	1	ACE_RABIT	ANGIOTENSIN-CONVERTING	5.50e+00
45	82	8.6	1108	1	CYGE_MOUSE	GUANYLYL CYCLASE GC-E	7.38e+00

ALIGNMENTS

RESULT	ID	SECR_PIG	STANDARD;	PRT;	131 AA.
AC	P01279;				
DT	21-JUL-1986 (REL. 01, CREATED)				
DT	01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)				
DT	01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)				
DE	SECRETIN PRECURSOR (FRAGMENT).				
GN	SCT.				
OS	SUS SCROFA (PIG), BOS TAURUS (BOVINE), AND CAVIA PORCELLUS (GUINEA PIG).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; EUTHERIA; ARTIODACTYLA.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	SPECIES=PIG;				
RX	MEDLINE; 90192795.				
RA	KOPIN A.S., WHEELER M.B., LEITER A.B.;				
RL	PROC. NATL. ACAD. SCI. U.S.A. 87:2299-2303(1990).				
RN	[2]				
RP	SEQUENCE OF 30-56.				
RC	SPECIES=PIG;				
RX	MEDLINE; 70282334.				
RA	MUTT V., JORPES J.E., MAGNUSSON S.;				
RL	EUR. J. BIOCHEM. 15:513-519(1970).				
RN	[3]				
RP	SEQUENCE OF 30-59 AND 92-131.				
RC	SPECIES=PIG;				
RX	MEDLINE; 90370867.				
RA	GAFFELIN G., JOERNVALL H., MUTT V.;				
RL	PROC. NATL. ACAD. SCI. U.S.A. 87:6781-6785(1990).				
RN	[4]				
RP	SYNTHESIS OF 30-131.				
RC	SPECIES=PIG;				
RA	BODANSZKY M., ONDETTI M.A., LEVINE S.D., NARAYANAN V.L., SALTZA M.V.,				
RL	SHEEHAN J.T., WILLIAMS N.J., SABO E.F.;				
RN	[5]				
RP	SEQUENCE OF 30-56.				
RC	SPECIES=BOVINE;				
RA	MEDLINE; 81237102.				
RL	CARLQUIST M., JOERNVALL H., MUTT V.;				
RN	[6]				
RP	SEQUENCE OF 30-56.				
RC	SPECIES=C PORCELLUS;				
RX	MEDLINE; 90254163.				
RA	BUSCAIL L., CAUVIN A., GOSSEN D., DE NEEF P., RATHE J.,				

```

FT REPEAT 112 144 ANAPHYLATOXIN-LIKE 3.
FT DOMAIN 176 215 EGF-LIKE 1.
FT DOMAIN 216 261 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 262 307 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 308 355 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 356 398 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 399 440 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 441 480 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 481 524 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 525 566 EGF-LIKE 9, CALCIUM-BINDING (INCOMPLETE).
FT DISULFID 36 61 BY SIMILARITY.
FT DISULFID 37 68 BY SIMILARITY.
FT DISULFID 50 69 BY SIMILARITY.
FT DISULFID 78 109 BY SIMILARITY.
FT DISULFID 91 110 BY SIMILARITY.
FT DISULFID 112 136 BY SIMILARITY.
FT DISULFID 113 143 BY SIMILARITY.
FT DISULFID 126 144 BY SIMILARITY.
FT DISULFID 180 190 BY SIMILARITY.
FT DISULFID 186 199 BY SIMILARITY.
FT DISULFID 201 214 BY SIMILARITY.
FT DISULFID 220 233 BY SIMILARITY.
FT DISULFID 227 242 BY SIMILARITY.
FT DISULFID 248 260 BY SIMILARITY.
FT DISULFID 266 279 BY SIMILARITY.
FT DISULFID 273 288 BY SIMILARITY.
FT DISULFID 306 325 BY SIMILARITY.
FT DISULFID 312 325 BY SIMILARITY.
FT DISULFID 319 334 BY SIMILARITY.
FT DISULFID 341 354 BY SIMILARITY.
FT DISULFID 360 373 BY SIMILARITY.
FT DISULFID 367 382 BY SIMILARITY.
FT DISULFID 384 397 BY SIMILARITY.
FT DISULFID 403 415 BY SIMILARITY.
FT DISULFID 411 424 BY SIMILARITY.
FT DISULFID 426 439 BY SIMILARITY.
FT DISULFID 445 454 BY SIMILARITY.
FT DISULFID 450 463 BY SIMILARITY.
FT DISULFID 465 479 BY SIMILARITY.
FT DISULFID 485 498 BY SIMILARITY.
FT DISULFID 494 507 BY SIMILARITY.
FT DISULFID 509 523 BY SIMILARITY.
FT CARBOHYD 98 98 POTENTIAL.
FT CARBOHYD 447 447 POTENTIAL.
FT CARBOHYD 535 535 POTENTIAL.
FT CARBOHYD 539 539 POTENTIAL.
FT CONFLICT 36 36 C -> S (IN REF. 2).
FT CONFLICT 41 42 HR -> SH (IN REF. 2).
FT SEQUENCE 566 AA; 61593 MW; A2D23E14 CRC32;

Query Match 10.1%; Score 96; DB 1; Length 566;
Best Local Similarity 57.1%; Pred.No. 9.43e-02;
Matches 16; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

Db 10 VPLPILGGTALLAAGVDADVLEACC 37
QY 13 VTLLELLLPALLSLGVDQPLPD-CC 39

RESULT 4
ID FBLN HUMAN STANDARD; PRT; 601 AA.
AC P23143;
DT 01-NOV-1991 (REL. 20, CREATED)
DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE FIBULIN-1, ISOFORM B PRECURSOR.
GN FBLN1
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91100426.

```

```

RA ARGRAVES W.S., TRAN H., BURGESS W.H., DICKERSON K.;
RL J. CELL BIOL. 111:3155-3164(1990).
RN [2]
RP SEQUENCE OF 30-44.
RX MEDLINE; 89354537.
RA ARGRAVES W.S., DICKERSON K., BURGESS W.H., RUOSLAHTI E.;
RL CELL 58:623-629(1989).
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX.
CC -!- ALTERNATIVE PRODUCTS: FOUR FORMS OF FIBULIN-1; A (AC P23142), B
(SHOWN HERE), C (AC P23144) AND D (AC P37888); DIFFERING ONLY IN
THEIR C-TERMINAL REGIONS, ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
CC EMBL; X53742; G31417;
DR PIR; A32826; A32826.
DR PIR; B36346; B36346.
DR MIM; 135820;
DR PROSITE; PS00010; ASX_HYDROXYL; 4.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; 3.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 3.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 8.
KW SIGNAL; ALTERNATIVE SPLICING; GLYCOPROTEIN; EXTRACELLULAR MATRIX;
KW REPEAT; PLASMA; EGF-LIKE DOMAIN; CALCIUM-BINDING.
FT SIGNAL 1 29
FT CHAIN 30 601 FIBULIN-1, ISOFORM B.
FT DOMAIN 36 144 3 X ANAPHYLATOXIN REPEATS.
FT REPEAT 77 76 ANAPHYLATOXIN-LIKE 1.
FT REPEAT 112 144 ANAPHYLATOXIN-LIKE 2.
FT REPEAT 112 144 ANAPHYLATOXIN-LIKE 3.
FT DOMAIN 176 215 EGF-LIKE 1.
FT DOMAIN 216 261 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 262 307 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 308 355 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 356 398 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 399 440 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 441 480 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 481 524 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 525 586 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT DISULFID 36 61 BY SIMILARITY.
FT DISULFID 37 68 BY SIMILARITY.
FT DISULFID 50 69 BY SIMILARITY.
FT DISULFID 78 109 BY SIMILARITY.
FT DISULFID 91 110 BY SIMILARITY.
FT DISULFID 112 136 BY SIMILARITY.
FT DISULFID 113 143 BY SIMILARITY.
FT DISULFID 126 144 BY SIMILARITY.
FT DISULFID 180 190 BY SIMILARITY.
FT DISULFID 186 199 BY SIMILARITY.
FT DISULFID 201 214 BY SIMILARITY.
FT DISULFID 220 233 BY SIMILARITY.
FT DISULFID 227 242 BY SIMILARITY.
FT DISULFID 248 260 BY SIMILARITY.
FT DISULFID 266 279 BY SIMILARITY.
FT DISULFID 273 288 BY SIMILARITY.
FT DISULFID 294 306 BY SIMILARITY.
FT DISULFID 312 325 BY SIMILARITY.
FT DISULFID 319 334 BY SIMILARITY.
FT DISULFID 341 354 BY SIMILARITY.
FT DISULFID 360 373 BY SIMILARITY.
FT DISULFID 367 382 BY SIMILARITY.
FT DISULFID 384 397 BY SIMILARITY.
FT DISULFID 403 415 BY SIMILARITY.
FT DISULFID 411 424 BY SIMILARITY.
FT DISULFID 426 439 BY SIMILARITY.
FT DISULFID 445 454 BY SIMILARITY.
FT DISULFID 450 463 BY SIMILARITY.
FT DISULFID 465 479 BY SIMILARITY.
FT DISULFID 485 498 BY SIMILARITY.
FT DISULFID 494 507 BY SIMILARITY.
FT DISULFID 509 523 BY SIMILARITY.
FT DISULFID 529 542 BY SIMILARITY.

```


Query Match 9.7%; Score 93; DB 1; Length 1061;
 Best Local Similarity 48.6%; Pred. No. 2.50e-01;
 Matches 17; Conservative 6; Mismatches 10; Indels 2; Gaps 2;

Db 1 MFGPRRPGASRLRLLLPLLLPLLLLRSHAGNL 35
 QY 3 LPSTKVPMAA-VTLLLLLLPLPALLSL-GVDAQPL 35

RESULT 9
 ID C10B_HUMAN STANDARD; PRT; 251 AA.
 AC P02746;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 01-JUL-1993 (REL. 26, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE COMPLEMENT C1Q SUBCOMPONENT, B CHAIN PRECURSOR.
 GN C10B.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 [1]
 SEQUENCE FROM N.A.
 RX MEDLINE; 86076906.
 RA REID K.B.M.;
 RL BIOCHEM. J. 231:729-735(1985).
 RN [2]
 RP SEQUENCE OF 26-133.
 RX MEDLINE; 80020137.
 RA REID K.B.M.;
 RL BIOCHEM. J. 179:367-371(1979).
 RN [3]
 RP SEQUENCE OF 26-193.
 RX MEDLINE; 79041552.
 RA REID K.B.M.; THOMPSON E.O.P.;
 RL BIOCHEM. J. 173:863-868(1978).
 RN [4]
 RP SEQUENCE OF 134-251.
 RX MEDLINE; 82283890.
 RA REID K.B.M.; GAGNON J.; FRAMPTON J.;
 RL BIOCHEM. J. 203:559-569(1982).
 RN [5]
 RP SEQUENCE OF 224-251 FROM N.A.
 RC TISSUE=LIVER;
 RX MEDLINE; 85038855.
 RA REID K.B.M.; BENTLEY D.R.; WOOD K.J.;
 RL PHLOS. TRANS. R. SOC. LOND., B, BIOL. SCI. 306:345-354(1984).
 CC -!- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD
 C1. THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
 COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
 C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
 TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
 FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
 CC -!- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q,
 R AND S IN THE MOLAR RATION OF 1:2:2.
 CC -!- SUBUNIT: C1Q SUBCOMPONENT IS COMPOSED OF NINE SUBUNITS, SIX OF
 WHICH ARE DISULFIDE-LINKED DIMERS OF THE A & B CHAINS, AND THREE
 OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE C CHAIN.
 CC -!- SIMILARITY: CONTAINS A C1Q DOMAIN.
 DR EMBL; X03084; G573114;
 DR EMBL; X36278; G179638;
 DR PIR; A03206; C1HQ0B.
 DR PIR; B23422; B23422.
 DR HSP; P19999; ICLG.
 DR MIM; 120570;
 DR PROSITE; PS01113; C1Q; 1.
 KW COMPLEMENT PATHWAY; PLASMA; HYDROXYLATION; GLYCOPROTEIN; COLLAGEN;
 REPEAT; SIGNAL.
 FT SIGNAL 1 25
 FT CHAIN 26 251
 FT DOMAIN 29 112
 FT DOMAIN 113 251
 FT MOD_RES 26 26
 FT DISULFID 29 29
 INTERCHAIN (WITH CYS-26 IN CHAIN A).

FT MOD_RES 33 33 HYDROXYLATION.
 FT MOD_RES 36 36 HYDROXYLATION.
 FT MOD_RES 39 39 HYDROXYLATION.
 FT MOD_RES 42 42 HYDROXYLATION.
 FT MOD_RES 51 51 HYDROXYLATION.
 FT MOD_RES 54 54 HYDROXYLATION.
 FT MOD_RES 57 57 HYDROXYLATION.
 FT CARBOHYD 57 57 GLUCOSYLALACTOSE.
 FT MOD_RES 60 60 HYDROXYLATION.
 FT CARBOHYD 60 60 GLUCOSYLALACTOSE.
 FT MOD_RES 63 63 HYDROXYLATION.
 FT MOD_RES 75 75 HYDROXYLATION.
 FT MOD_RES 81 81 HYDROXYLATION.
 FT MOD_RES 84 84 HYDROXYLATION.
 FT MOD_RES 90 90 HYDROXYLATION.
 FT MOD_RES 96 96 HYDROXYLATION.
 FT CARBOHYD 96 96 GLUCOSYLALACTOSE.
 FT MOD_RES 99 99 HYDROXYLATION.
 FT MOD_RES 102 102 HYDROXYLATION.
 FT MOD_RES 105 105 HYDROXYLATION.
 FT MOD_RES 108 108 HYDROXYLATION.
 FT CARBOHYD 108 108 GLUCOSYLALACTOSE.
 FT VARIANT 174 174 S -> Z (POTENTIAL).
 FT CONFLICT 26 26 Q -> E (IN REF. 2).
 FT CONFLICT 83 83 N -> D (IN REF. 2).
 FT CONFLICT 98 98 G -> P (IN REF. 2 AND 3).
 SQ SEQUENCE 251 AA; 26459 MW; 23D5197F CRC32;

Query Match 9.5%; Score 91; DB 1; Length 251;
 Best Local Similarity 53.3%; Pred. No. 4.74e-01;
 Matches 8; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Db 2 KIPWGPVLMILL 16
 QY 7 KVPWAAVTLMLLL 21

RESULT 10
 ID LCAT_RAT STANDARD; PRT; 440 AA.
 AC P18424;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)
 DE PHOSPHATIDYLCHOLINE-STEROL ACYLTRANSFERASE PRECURSOR (EC 2.3.1.43)
 DE (LECITHIN-CHOLESTEROL ACYLTRANSFERASE) (PHOSPHOLIPID-CHOLESTEROL
 ACYLTRANSFERASE).
 GN LCAT.
 OS RATTUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 90384859.
 RA MERONI G.; MARGARETTI N.; MAGNAGHI P.; TARAMELLI R.;
 RL NUCLEIC ACIDS RES. 18:5308-5308(1990).
 CC -!- FUNCTION: CENTRAL ENZYME IN THE EXTRACELLULAR METABOLISM OF PLASMA
 LIPOPROTEINS. AMONG OTHER SUBSTRATES IT ESTERIFIES THE FREE
 CHOLESTEROL TRANSPORTED IN PLASMA LIPOPROTEINS.
 CC -!- CATALYTIC ACTIVITY: PHOSPHATIDYLCHOLINE + STEROL -> STEROL ESTER +
 1-ACYLGLYCEROPHOSPHOCHOLINE (PALMITOYL, OLEOYL, AND LINOLEOYL CAN
 BE TRANSFERRED; A NUMBER OF STEROLS, INCLUDING CHOLESTEROL, CAN
 ACT AS ACCEPTOR).
 CC -!- ENZYME REGULATION: APOLIPOPROTEIN A-I IS A POTENT ACTIVATOR FOR
 THIS ENZYME.
 CC -!- SIMILARITY: PARTIAL WITH LIPASES (PANCREATIC, GASTRIC, HEPATIC,
 LINGUAL, LIPOPROTEIN, BACTERIAL, ETC.).
 DR EMBL; X54096; G56564;
 DR PIR; S11214; XXRN.
 DR PROSITE; PS00120; LIPASE_SER; 1.
 KW TRANSFERASE; ACYLTRANSFERASE; LIPID METABOLISM; GLYCOPROTEIN; SIGNAL.
 FT SIGNAL 1 24
 FT CHAIN 25 440
 FT ACT_SITE 205 205 LECITHIN-CHOLESTEROL ACYLTRANSFERASE.
 FT ACT_SITE 205 205 CHARGE RELAY SYSTEM (BY SIMILARITY).

DR EMBL; M26268; G187025; -
 DR EMBL; X04981; G34287; -
 DR EMBL; M17959; G38658; -
 DR PIR; A00571; XXHUN.
 DR PIR; A29661; A29661.
 DR PIR; A25575; A25575.
 DR PIR; J00036; J00036.
 DR MIM; 136120; -
 DR MIM; 245900; -
 DR PROSITE; PS00120; LIPASE_SER; 1.
 KW TRANSFERASE; ACYLTRANSFERASE; LIPID METABOLISM; GLYCOPROTEIN; SIGNAL;
 KW POLYMORPHISM; DISEASE MUTATION.
 FT SIGNAL 1 24
 FT CHAIN 25 440
 FT ACT_SITE 205 205
 FT DISULFID 74 98
 FT FT 337 380
 FT CARBOHYD 44 44
 FT CARBOHYD 108 108
 FT CARBOHYD 296 296
 FT CARBOHYD 408 408
 FT CARBOHYD 431 431
 FT CARBOHYD 433 433
 FT VARIANT 17 17
 FT VARIANT 34 34
 FT VARIANT 54 54
 FT VARIANT 57 57
 FT VARIANT 117 117
 FT VARIANT 147 147
 FT VARIANT 159 159
 FT VARIANT 164 164
 FT VARIANT 170 170
 FT VARIANT 180 180
 FT VARIANT 182 182
 FT VARIANT 233 233
 FT VARIANT 252 252
 FT VARIANT 276 276
 FT VARIANT 317 317
 FT VARIANT 345 345
 FT VARIANT 371 371
 FT CONFLICT 257 257
 FT CONFLICT I -> H (IN REF. 3).
 SQ SEQUENCE 440 AA; 49578 MW; 69727CDF CRC32;
 Query Match 9.5%; Score 91; DB 1; Length 440;
 Best Local Similarity 81.3%; Pred. No. 4.74e-01;
 Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 7 PWQWVILLGLLLPPA 22
 || ||||| |||||
 9 PWAAVTILLGLLLPPA 24
 RESULT 12
 ID PDGB_HUMAN STANDARD; PRT; 241 AA.
 AC P01127;
 DT 21-JUL-1986 (REL. 01, CREATED)
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE PLATELET-DERIVED GROWTH FACTOR, B CHAIN PRECURSOR (PDGF B-CHAIN)
 DE (PDGF-2) (BACAPLERMIN).
 GN PDGB OR C-SIS.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 84250225.
 RA JOSEPHS S.F., RATNER L., CLARKE M.F., WESTIN E.H., REITZ M.S.,
 RA WONG-STAAAL F.;
 RL SCIENCE 225:636-639(1984).
 RN [2]

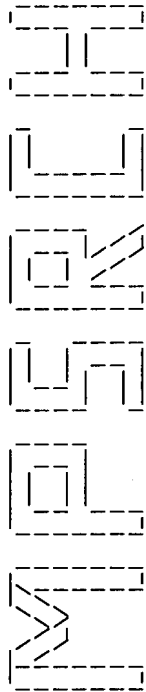
RP SEQUENCE FROM N.A.
 RX MEDLINE; 86205961.
 RA RAO C.D., IGARASHI H., CHIU I.-M., ROBBINS K.C., AARONSON S.A.;
 RL PROC. NATL. ACAD. SCI. U.S.A. 83:2392-2396(1986).
 RN [3]
 RP SEQUENCE OF 22-241 FROM N.A.
 RX MEDLINE; 84205633.
 RA CHIU I.-M., REDDY E.P., GIVOL D., ROBBINS K.C., TRONICK S.R.,
 RA AARONSON S.A.;
 RL CELL 37:123-129(1984).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 85296313.
 RA COLLINS T., GINSBURG D., BOSS J.M., ORKIN S.H., POBER J.S.;
 RL NATURE 316:748-750(1985).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 85269623.
 RA RATNER L., JOSEPHS S.F., JARRETT R., REITZ M.S., WONG-STAAAL F.;
 RL NUCLEIC ACIDS RES. 13:5007-5018(1985).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 87217119.
 RA RAO C.D., IGARASHI H., PECH M.W., ROBBINS K.C., AARONSON S.A.;
 RL COLD SPRING HARB. SYMP. QUANT. BIOL. 51:959-966(1986).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX BURNESS J., ODELL C.;
 RL SUBMITTED (OCT-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [8]
 RP SEQUENCE OF 26-241 FROM N.A.
 RX MEDLINE; 86164981.
 RA WEICH H.A., SEBALD W., SCHAIRER H.U., HOPPE J.;
 RL FEBS LETT. 198:344-348(1986).
 RN [9]
 RP SEQUENCE OF 82-110.
 RX MEDLINE; 83197379.
 RA ANTONIADES H.N., HUNKAPILLER M.W.;
 RL SCIENCE 220:963-965(1983).
 RN [10]
 RP SEQUENCE OF 82-112.
 RX MEDLINE; 83244981.
 RA WATERFIELD M.D., SCRACE G.T., WHITTLE N., STROOBANT P., JOHNSON A.,
 RA WATSON A., WESTERMARK B., HELDIN C.H., HUANG J.S., DEUEL T.F.;
 RL NATURE 304:35-39(1983).
 RN [11]
 RP MUTAGENESIS. IMPORTANCE OF ARG-108 AND ILE-111 FOR RECEPTOR-BINDING.
 RX MEDLINE; 92097530.
 RA CLEMENTS J.M., BARDEN L.J., BLOXIDGE R.E., CATLIN G., COOK A.L.,
 RA CRAIG S., DRUMMOND A.H., EDWARDS R.M., FALLON A., GREEN D.R.,
 RA HELLEWELL P.G., KIRWIN P.M., NAYEE P.D., RICHARDSON S.J., BROWN D.,
 RA CHAHWALA S.B., SNAREY M., WINSLOW D.;
 RL EMBO J. 10:4113-4120(1991).
 RN [12]
 RP INTERCHAIN DISULFIDE BONDS.
 RX MEDLINE; 92283833.
 RA ANDERSSON M., OESTMAN A., BAECTROEM G., HELLMAN U.,
 RA GEORGE-NASCIMENTO C., WESTERMARK B., HELDIN C.-H.;
 RL J. BIOL. CHEM. 267:11260-11266(1992).
 RN [13]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RX MEDLINE; 93010987.
 RA OEFNER C., D'ARCY A., WINKLER F.K., EGGIMANN B., HOSANG M.;
 RL EMBO J. 11:3921-3926(1992).
 CC -!- FUNCTION: PLATELET-DERIVED GROWTH FACTOR IS A POTENT MITOGEN FOR
 CC CELLS OF MESENCHYMAL ORIGIN. BINDING OF THIS GROWTH FACTOR TO ITS
 CC AFFINITY RECEPTOR ELICITS A VARIETY OF CELLULAR RESPONSES. IT IS
 CC RELEASED BY PLATELETS UPON WOUNDING AND PLAYS AN IMPORTANT ROLE
 CC IN STIMULATING ADJACENT CELLS TO GROW AND THEREBY HEAL THE WOUND.
 CC -!- SUBUNIT: ANTIPARALLEL DISULFIDE-LINKED DIMER OF NONIDENTICAL (A
 CC AND B) CHAINS. HOMODIMERS OF A AND B CHAINS ARE IMPLICATED IN
 CC TRANSFORMATION PROCESSES.
 CC -!- A-A AND B-B, AS WELL AS A-B, DIMERS CAN BIND TO THE PDGF RECEPTOR.

Db 121 HPPSRFRLPRLRLRYTAHLA-RLRL-RRAGGGAPEPPATPATPATPATPAT 178
QY 68 RRGPP-GLQRLRLQNGHAGILTMGRAGAELEFPYCPGRCPTATATAPRG 126
Db 179 PARV 182
QY 127 GSRV 130

RESULT 15
ID 41BL_HUMAN STANDARD; PRT; 254 AA.
AC P41273;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE 4-1BB LIGAND (4-1BBU).
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
PM [1]
SEQUENCE FROM N.A.
MEDLINE: 94374434.
RA ALDERSON M.R., SMITH C.A., TOUGH T.W., DAVIS-SMITH T., ARMITAGE R.J.,
RA FALK B., ROUX E., BAKER E., SUTHERLAND G.R., DIN W.S., GOODWIN R.G.;
RL EUR. J. IMMUNOL. 24:2219-2227(1994).
CC -!- FUNCTION: INDUCES THE PROLIFERATION OF ACTIVATED PERIPHERAL BLOOD
CC T CELLS. MAY HAVE A ROLE IN ACTIVATION-INDUCED CELL DEATH (AICD).
CC MAY PLAY A ROLE IN COGNATE INTERACTIONS BETWEEN T CELLS AND
CC B CELLS/MACROPHAGES.
CC -!- SUBUNIT: HOMOTRIMER (POTENTIAL).
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, PLACENTA, LUNG, SKELETAL
CC MUSCLE AND KIDNEY.
CC -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL; U03398; G571323; -.
DR PROSITE; PS00251; TNF 1; 1.
DR PROSITE; PS00049; TNF 2; 1.
KW CYTOKINE; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT DOMAIN 1 28 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 29 49 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 50 254 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 35 41 POLY-LEU.
SQ SEQUENCE 254 AA; 26624 MW; C68C1B27 CRC32;

Query Match 9.28; Score 88; DB 1; Length 254;
Best Local Similarity 66.78; Pred. No. 1.21e+00;
Matches 12; Conservative 4; Mismatches 1; Indels 1; Gaps 1;
QY 27 LPWALVAGLLLLLLLLAA 44
8 VPWAAVT-LLLLLLPPA 24

Search completed: Thu Jul 30 09:17:05 1998
Job time : 14 secs.



(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:16:01 1998; MasPar time 7.56 Seconds
628.387 Million cell updates/sec
ular output not generated.

Title: >US-08-938-548A-6
Description: (1-130) from US08938548A.pep
Perfect Score: 954
Sequence: 1 MNLPSKVPWNAVILLILL.....GRCPTATATAPRGSSRV 130

Scoring table: PAM 150
Gap 11
Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir56
1:pir1 2:pir2 3:pir3 4:pir4 5:nrl3d

Statistics: Mean 39.951; Variance 87.022; scale 0.459

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description	Pred. No.
1	101	10.6	131	1	SEPG secretin precursor -	1.14e-01
2	100	10.5	495	2	catalase (EC 1.11.1.6	1.52e-01
3	98	10.3	641	2	probable intercellula	2.72e-01
4	96	10.1	438	1	phosphatidylcholine--	4.82e-01
5	96	10.1	601	2	fibulin 1 precursor,	4.82e-01
6	96	10.1	683	2	collagen alpha chain	4.82e-01
7	93	9.7	1027	2	matricellular peptide r	1.12e+00
8	93	9.7	1061	1	CYHVAR complement subcompone	1.12e+00
9	91	9.5	253	1	phosphatidylcholine--	1.96e+00
10	91	9.5	440	1	XXRHN phosphatidylcholine--	1.96e+00
11	91	9.5	440	1	XXRHN phosphatidylcholine--	1.96e+00
12	91	9.5	602	2	S39782 cyclooxygenase 1 - ra	1.96e+00
13	89	9.3	185	2	S58383 hypothetical protein	3.38e+00
14	89	9.3	230	2	A55030 platelet-derived grow	3.38e+00
15	89	9.3	241	1	PFHUG2 platelet-derived grow	3.38e+00
16	89	9.3	321	1	salivary proline-rich	3.38e+00
17	89	9.3	602	2	S69198 prostaglandin G/H syn	3.38e+00
18	88	9.2	245	2	FLT3/FLK2 ligand (clo	4.43e+00
19	88	9.2	491	2	JC6197 stromelysin 3 (EC 3.4	4.43e+00
20	88	9.2	492	2	A37445 phosphatidylinositol-	4.43e+00
21	87	9.1	226	1	TMVSS PDGF-related transfer	5.79e+00
22	87	9.1	271	2	A25659 PDGF-related transfer	5.79e+00
23	87	9.1	583	2	I50518 DNA binding protein E	5.79e+00

24	87	9.1	599	2	A36746 prostaglandin-endoper	5.79e+00
25	86	9.0	322	2	G6151 hypothetical protein	7.56e+00
26	86	9.0	334	2	S16296 ferric enterobactin t	7.56e+00
27	86	9.0	440	2	JC1502 phosphatidylcholine--	7.56e+00
28	86	9.0	1102	2	JH0717 guanylate cyclase (EC	7.56e+00
29	86	9.0	1663	1	C3RT complement C3 precurs	7.56e+00
30	85	8.9	236	2	A37930 placental lactogen pr	9.85e+00
31	85	8.9	241	1	PFMSGH platelet-derived grow	9.85e+00
32	85	8.9	255	2	S12255 hypothetical protein	9.85e+00
33	85	8.9	317	2	S28225 triacylglycerol lipas	9.85e+00
34	85	8.9	317	2	S57275 triacylglycerol lipas	9.85e+00
35	85	8.9	613	2	JC5501 endothelin receptor t	9.85e+00
36	84	8.8	91	2	J50455 gene E protein - phag	1.28e+01
37	84	8.8	102	2	S26409 protein 108 precursor	1.28e+01
38	84	8.8	315	2	S14276 triacylglycerol lipas	1.28e+01
39	84	8.8	329	2	D41344 lutropin-choriogonado	1.28e+01
40	84	8.8	331	2	C41344 lutropin-choriogonado	1.28e+01
41	84	8.8	342	2	S68129 NADH dehydrogenase (u	1.28e+01
42	84	8.8	383	2	B48122 GTPase-activating pro	1.28e+01
43	84	8.8	384	2	A48122 GTPase-activating pro	1.28e+01
44	84	8.8	695	2	A41344 lutropin-choriogonado	1.28e+01
45	84	8.8	701	2	S61239 hypothetical protein	1.28e+01

ALIGNMENTS

RESULT ENTRY	1	SEPG	#type complete
TITLE		secretin precursor - pig	
ORGANISM		#formal_name Sus scrofa domestica	#common_name domestic pig
DATE		24-Apr-1984	#sequence_revision 12-Apr-1996
			#text_change
ACCESSIONS		B35094; A01544; A36052	
REFERENCE		A35094	
#authors		Kopin, A.S.; Wheeler, M.B.; Leiter, A.B.	
#journal		Proc. Natl. Acad. Sci. U.S.A. (1990)	87:2299-2303
#title		Secretin: structure of the precursor and tissue distribution of the mRNA.	
#cross-references		MUID:90192795	
#accession		B35094	
#molecule_type		mrna	
#residues		1-131	#label KOP
#cross-references		GB:M31496; NID:g164670; PID:g164671	

REFERENCE		A91147	
#authors		Mutt, V.; Jorpes, J.E.; Magnusson, S.	
#journal		Eur. J. Biochem. (1970)	15:513-519
#title		Structure of porcine secretin. The amino acid sequence.	
#cross-references		MUID:70282334	
#accession		A01544	
#molecule_type		protein	
#residues		30-56	#label MUT
#note		tryptic peptides were sequenced	

REFERENCE		A36052	
#authors		Gafvelin, G.; Joernvall, H.; Mutt, V.	
#journal		Proc. Natl. Acad. Sci. U.S.A. (1990)	87:6781-6785
#title		Processing of prosecretin: isolation of a secretin precursor from porcine intestine.	
#cross-references		MUID:90370867	
#accession		A36052	
#status		preliminary	
#molecule_type		protein	
#residues		30-59, 'R', 92-131	#label GAF

REFERENCE		A90916	
#authors		Bodanszky, M.; Ondetti, M.A.; Levine, S.D.; Narayanan, V.L.; Salza, M.V.; Sheehan, J.T.; Williams, N.J.; Sabo, E.F.	
#journal		Chem. Ind. (1966)	:1757-1758
#title		Synthesis of a heptacosapeptide amide with the hormonal activity of secretin.	
#contents		synthesis confirmed the proposed structure of the natural hormone	
#note			
CLASSIFICATION		#superfamily glucagon	
KEYWORDS		amidated carboxyl end; duodenal mucosa; duplication; hormone;	

```
ORGANISM #formal_name Homo sapiens #common_name man
DATE 19-Apr-1991 #sequence_revision 19-Apr-1991 #text_change
29-Aug-1997
ACCESSIONS B36346
REFERENCE B36346
#authors Argraves, W.S.; Tran, H.; Burgess, W.H.; Dickerson, K.
#journal J. Cell Biol. (1990) 111:3155-3164
#title Fibulin is an extracellular matrix and plasma glycoprotein
with repeated domain structure.
#cross-references MUID:91100426
#accession B36346
#molecule_type mRNA
#residues 1-601 #label ARG
#cross-references GB:X53742
GENETICS
#gene GDB:FBLN1; FBLN
#cross-references GDB:278285; OMIM:135820
#map_position 22q13.3-22q13.3
KEYWORDS alternative splicing
SUMMARY #length 601 #molecular-weight 65485 #checksum 6896
Query Match 10.1%; Score 96; DB 2; Length 601;
Best Local Similarity 57.1%; Pred. No. 4.82e-01;
Matches 16; Conservative 3; Mismatches 8; Indels 1; Gaps 1;
Db 10 VPLPILLLGLALLAAGVDADVLEACC 37
QY 13 VTLLLLPPALLSLGVDQPLPD-CC 39
#cross-references MUID:91100426
#accession B36346
#molecule_type mRNA
#residues 1-683 #label ARG
#cross-references GB:X53743; NID:g31418; PID:g31419
ACCESSIONS A36346
REFERENCE A36346
#authors Argraves, W.S.; Tran, H.; Burgess, W.H.; Dickerson, K.
#journal J. Cell Biol. (1990) 111:3155-3164
#title Fibulin is an extracellular matrix and plasma glycoprotein
with repeated domain structure.
#cross-references MUID:91100426
#accession C36346
#molecule_type mRNA
#residues 1-683 #label ARG
#cross-references GB:X53743; NID:g31418; PID:g31419
ACCESSIONS A36346
REFERENCE A36346
#authors Argraves, W.S.; Dickerson, K.; Burgess, W.H.; Ruoslahti, E.
#journal Cell (1989) 58:623-629
#title Fibulin, a novel protein that interacts with the fibronectin
receptor beta-subunit cytoplasmic domain.
#cross-references MUID:89354537
#accession A32826
#molecule_type protein
#residues 30-35, 'SX', 38-40, 'SH', 43-44 #label AR3
GENETICS
#gene GDB:FBLN1; FBLN
#cross-references GDB:278285; OMIM:135820
#map_position 22q13.3-22q13.3
KEYWORDS alternative splicing; glycoprotein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-683 #product fibulin 1 splice form C #status predicted
#label MAT\
98,535,539 #binding_site carbohydrate (Asn) (covalent) #status
predicted
```

```
SUMMARY #length 683 #molecular-weight 74475 #checksum 7443
Query Match 10.1%; Score 96; DB 2; Length 683;
Best Local Similarity 57.1%; Pred. No. 4.82e-01;
Matches 16; Conservative 3; Mismatches 8; Indels 1; Gaps 1;
Db 10 VPLPILLLGLALLAAGVDADVLEACC 37
QY 13 VTLLLLPPALLSLGVDQPLPD-CC 39
#cross-references MUID:91100426
#accession B36346
#molecule_type mRNA
#residues 1-601 #label ARG
#cross-references GB:X53742
GENETICS
#gene GDB:FBLN1; FBLN
#cross-references GDB:278285; OMIM:135820
#map_position 22q13.3-22q13.3
KEYWORDS alternative splicing
SUMMARY #length 601 #molecular-weight 65485 #checksum 6896
Query Match 10.1%; Score 96; DB 2; Length 601;
Best Local Similarity 57.1%; Pred. No. 4.82e-01;
Matches 16; Conservative 3; Mismatches 8; Indels 1; Gaps 1;
Db 10 VPLPILLLGLALLAAGVDADVLEACC 37
QY 13 VTLLLLPPALLSLGVDQPLPD-CC 39
#cross-references MUID:91100426
#accession B36346
#molecule_type mRNA
#residues 1-683 #label ARG
#cross-references GB:X53743; NID:g31418; PID:g31419
ACCESSIONS A36346
REFERENCE A36346
#authors Argraves, W.S.; Tran, H.; Burgess, W.H.; Dickerson, K.
#journal J. Cell Biol. (1990) 111:3155-3164
#title Fibulin is an extracellular matrix and plasma glycoprotein
with repeated domain structure.
#cross-references MUID:91100426
#accession C36346
#molecule_type mRNA
#residues 1-683 #label ARG
#cross-references GB:X53743; NID:g31418; PID:g31419
ACCESSIONS A36346
REFERENCE A36346
#authors Argraves, W.S.; Dickerson, K.; Burgess, W.H.; Ruoslahti, E.
#journal Cell (1989) 58:623-629
#title Fibulin, a novel protein that interacts with the fibronectin
receptor beta-subunit cytoplasmic domain.
#cross-references MUID:89354537
#accession A32826
#molecule_type protein
#residues 30-35, 'SX', 38-40, 'SH', 43-44 #label AR3
GENETICS
#gene GDB:FBLN1; FBLN
#cross-references GDB:278285; OMIM:135820
#map_position 22q13.3-22q13.3
KEYWORDS alternative splicing; glycoprotein
FEATURE
1-29 #domain signal sequence #status predicted #label SIG\
30-683 #product fibulin 1 splice form C #status predicted
#label MAT\
98,535,539 #binding_site carbohydrate (Asn) (covalent) #status
predicted
```

```

#accession A23422
#molecule_type mRNA
#residues 28-253 ##label RE1
##cross-references EMBL:X03084
##note the authors translated the codon ACA for residue 46 as Ile
REFERENCE
#authors Reid, K.B.M.
#journal Blochem. J. (1979) 179:367-371
#title Complete amino acid sequences of the three collagen-like regions present in subcomponent Clq of the first component of human complement.
#cross-references MUID:80020137
#accession B90304
#molecule_type protein
#residues 'E',29-84,'D',86-99,'P',101-135 ##label RE5
REFERENCE
#authors Reid, K.B.M.; Thompson, E.O.P.
#journal Blochem. J. (1978) 173:863-868
#title Amino acid sequence of the N-terminal 108 amino acid residues of the B chain of subcomponent Clq of the first component of human complement.
#cross-references MUID:79041552
#accession A90301
#molecule_type protein
#residues 28-99,'P',101-195 ##label RE3
REFERENCE
#authors Reid, K.B.M.; Gagnon, J.; Frampton, J.
#journal Blochem. J. (1982) 203:559-569
#title Completion of the amino acid sequences of the A and B chains of subcomponent Clq of the first component of human complement.
#cross-references MUID:82283890
#accession B90315
#molecule_type protein
#residues 136-253 ##label RE4
#note 176-Glx may also be present
COMMENT The first component of complement is a calcium-dependent complex of the three subcomponents Clq, Clr, and C1s. Subcomponent Clq binds to immunoglobulin complexes, with resulting serial activation of C1r (enzyme), C1s (proenzyme), and the other eight components of complement.
The Clq subcomponent is composed of nine subunits, six of which are disulfide-linked dimers of the A (see C1HUQA) and B chains, and three of which are disulfide-linked dimers of the C (see C1HUQC) chain. Equimolar amounts of the A, B, and C chains are found after reduction of the disulfide bonds.
GENETICS
#gene GDB:CIQB
##cross-references GDB:119043; OMIM:120570
#map_position lp36.3-lp34.1
CLASSIFICATION #superfamily complement subcomponent Clq chain A; complement Clq carboxyl-terminal homology
KEYWORDS complement pathway; glycoprotein; heterodimer; hydroxylysine; hydroxyproline; plasma; pyroglutamic acid; triple helix
FEATURE
1-27 #domain signal sequence #status predicted #label SIG\
28-253 #product complement subcomponent Clq chain B #status experimental #label MAT\
33-116 #domain collagenous, triple helix #label COL\
123-249 #domain complement Clq carboxyl-terminal homology #label Clq\
28 #modified_site pyrrolidone carboxylic acid (Gln) (in mature form) #status experimental\
31 #disulfide_bonds interchain (to chain A-26) #status experimental\
35,38,41,53,56,65, #modified_site 4-hydroxyproline (Pro) #status experimental\
83,86,101,104,107 #modified_site 5-hydroxylysine (Lys) #status experimental\
59,62,77,92,98, #binding_site carboxylate (Lys) (covalent) #status experimental\
110
59,62,98,110

```

```

SUMMARY #length 253 #molecular-weight 26722 #checksum 7399
Query Match 9.5%; Score 91; DB 1; Length 253;
Best Local Similarity 53.3%; Pred. No. 1.96e+00;
Matches 8; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Db 4 KIPWGSIPVLMILL 18
|:|:::|:|::|
QY 7 KVPWAAVTLELLLL 21

RESULT 10 XXRTN #type complete
ENTRY phosphatidylcholine--sterol O-acyltransferase (EC 2.3.1.43)
TITLE precursor - rat
ALTERNATE_NAMES lecithin--cholesterol acyltransferase precursor;
phospholipid--cholesterol acyltransferase precursor
ORGANISM #formal_name Rattus norvegicus #common_name Norway rat
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Sep-1997
ACCESSIONS S11214; S11302
REFERENCE S11214
#authors Meroni, G.; Malgaretti, N.; Magnaghi, P.; Taramelli, R.
#journal Nucleic Acids Res. (1990) 18:5308
#title Nucleotide sequence of the cDNA for lecithin-cholesterol acyltransferase (LCAT) from the rat.
#cross-references MUID:90384859
#accession S11214
#molecule_type mRNA
#residues 1-440 ##label MER
##cross-references EMBL:X54096
REFERENCE S11302
#authors Taramelli, R.
#submission submitted to the EMBL Data Library, July 1990
#accession S11302
#molecule_type mRNA
#residues 1-389,'G',391-440 ##label TAR
##cross-references EMBL:X54096; NID:G56563; PID:G56564
COMMENT The active enzyme catalyzes the transfer of acyl groups from lecithin to sterol to form sterol esters. Palmitoyl, oleoyl, and linoleoyl residues can be transferred; a number of sterols, including cholesterol, can act as acceptor. Apolipoprotein A-I is a potent activator for this enzyme.
GENETICS
#gene LCAT
CLASSIFICATION #superfamily phosphatidylcholine--sterol acyltransferase
KEYWORDS acyltransferase; glycoprotein; lipid metabolism; lipoprotein
FEATURE
1-24 #domain signal sequence #status predicted #label SIG\
25-438 #product phosphatidylcholine--sterol acyltransferase #status predicted #label MAT\
44,108,296,408 #binding_site carboxylate (Asn) (covalent) #status predicted
SUMMARY #length 440 #molecular-weight 49741 #checksum 6639
Query Match 9.5%; Score 91; DB 1; Length 440;
Best Local Similarity 62.5%; Pred. No. 1.96e+00;
Matches 15; Conservative 2; Mismatches 5; Indels 2; Gaps 1;

Db 1 MGLPGS--PMQWVLLGLLLPPA 22
|:|::|:|:|:|:|:|:|
QY 1 MNLPSKVPWAAVTLELLLLPPA 24

RESULT 11 XXHUN #type complete
ENTRY phosphatidylcholine--sterol O-acyltransferase (EC 2.3.1.43)
TITLE precursor - human
ALTERNATE_NAMES lecithin--cholesterol acyltransferase precursor;
phospholipid--cholesterol acyltransferase precursor
ORGANISM #formal_name Homo sapiens #common_name man
DATE 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change

```

Db 1 MSRRSLQFPLLLLLLPPLPPVLLTDAGVSPVIP-CC 39
 QY 1 MMLPSTKVPWAAVTLTLLLLLPALLS-LGVDAQPLPDC 39

RESULT 13

ENTRY S58383 #type complete
 TITLE hypothetical protein 2 - human
 ORGANISM #formal_name Homo sapiens #common_name man
 DATE 12-Feb-1998 #sequence_revision 20-Feb-1998 #text_change 20-Feb-1998

ACCESSIONS S58383
 REFERENCE S58382

#authors Dirks, R.P.H.; Onnekink, C.; Jansen, H.J.; de Jong, A.; Bloemers, H.P.J.
 #journal Nucleic Acids Res. (1995) 23:2815-2822
 #title A novel human c-sis mRNA species is transcribed from a promoter in c-sis intron 1 and contains the code for an alternative PDGF B-like protein.

#accession S58383
 #status preliminary
 #molecule_type mRNA
 #residues 1-185 #label DIR
 #cross-references EMBL:X83705
 SUMMARY #length 185 #molecular-weight 20774 #checksum 2728

Query Match 9.3%; Score 89; DB 2; Length 185;
 Best Local Similarity 43.6%; Pred. No. 3.38e+00;
 Matches 17; Conservative 11; Mismatches 6; Indels 5; Gaps 5;

Db 29 LQRLHGDGEGDGAELDNMTSRSHGGLESIA-RGRR 66

QY 79 LQRLQAN-GNH-AAGI-LTMGR-RAGALEPYPGRR 113

RESULT 14

ENTRY A55030 #type fragment
 TITLE platelet-derived growth factor chain B precursor - human (fragment)

ORGANISM #formal_name Homo sapiens #common_name man
 DATE 18-Nov-1994 #sequence_revision 18-Nov-1994 #text_change 03-May-1996

ACCESSIONS A55030
 REFERENCE A55030
 #authors Johnsson, A.; Heldin, C.H.; Wasteson, A.; Westermark, B.; Duell, T.F.; Huang, J.S.; Seeburg, P.H.; Gray, A.; Ullrich, A.; Scraze, G.; Stroobant, P.; Waterfield, M.D.

#journal EMBO J. (1984) 3:921-928
 #title The c-sis gene encodes a precursor of the B chain of platelet-derived growth factor.

#accession A55030
 #status preliminary
 #molecule_type DNA
 #residues 1-230 #label JOH
 #cross-references GB:X00556; GB:X00559; GB:X00560; GB:X00561; GB:X00562
 CLASSIFICATION #superfamily platelet-derived growth factor
 SUMMARY #length 230 #checksum 3580

Query Match 9.3%; Score 89; DB 2; Length 230;
 Best Local Similarity 43.6%; Pred. No. 3.38e+00;
 Matches 17; Conservative 11; Mismatches 6; Indels 5; Gaps 5;

Db 33 LQRLHGDGEGDGAELDNMTSRSHGGLESIA-RGRR 70

QY 79 LQRLQAN-GNH-AAGI-LTMGR-RAGALEPYPGRR 113

RESULT 15

ENTRY PFHUG2 #type complete
 TITLE platelet-derived growth factor chain B precursor - human
 ALTERNATE_NAMES PDGF B-chain; PDGF-B; PDGF-II; PDGF-related transforming protein (sis)

ORGANISM #formal_name Homo sapiens #common_name man
 DATE 18-Apr-1984 #sequence_revision 20-Sep-1984 #text_change 20-Mar-1998
 ACCESSIONS A94276; A21024; A23532; A93366; A25141; A94271; A93308;
 REFERENCE A43499; S56115; I57635; I37266; A01380; A94622
 #authors Josephs, S.F.; Ratner, L.; Clarke, M.F.; Westin, E.H.; Reitz, M.S.; Wong-Staal, F.

#journal Science (1984) 225:636-639
 #title Transforming potential of human c-sis nucleotide sequences encoding platelet-derived growth factor.

#cross-references MUID:84350225
 #accession A94276

#molecule_type DNA
 #residues 1-241 #label JOS
 #cross-references GB:K01401; NID:g338206; PID:g338209

REFERENCE A21024
 #authors Chiu, I.M.; Reddy, E.P.; Givol, D.; Robbins, K.C.; Tronick, S.R.; Aaronson, S.A.

#journal Cell (1984) 37:123-129
 #title Nucleotide sequence analysis identifies the human c-sis proto-oncogene as a structural gene for platelet-derived growth factor.

#cross-references MUID:84205633
 #accession A21024

#molecule_type DNA
 #residues 17-20, 'RQ', 22-241 #label CH2
 REFERENCE A23532

#authors Rao, C.D.; Igarashi, H.; Chiu, I.M.; Robbins, K.C.; Aaronson, S.A.

#journal Proc. Natl. Acad. Sci. U.S.A. (1986) 83:2392-2396
 #title Structure and sequence of the human c-sis/platelet-derived growth factor 2 (SIS/PDGF2) transcriptional unit.

#cross-references MUID:86205961
 #accession A23532

#molecule_type mRNA
 #residues 1-241 #label RAO
 #cross-references GB:M12783; GB:M16288; NID:g338210; PID:g338211

REFERENCE A93366
 #authors Collins, T.; Ginsburg, D.; Boss, J.M.; Orkin, S.H.; Pober, J.S.

#journal Nature (1985) 316:748-750
 #title Cultured human endothelial cells express platelet-derived growth factor B chain: cDNA cloning and structural analysis.

#cross-references MUID:85296313
 #accession A93366

#molecule_type mRNA
 #residues 1-241 #label COL
 REFERENCE A25141

#authors Weich, H.A.; Sebal, W.; Schairer, H.U.; Hoppe, J.
 #journal FEBS Lett. (1986) 198:344-348
 #title The human osteosarcoma cell line U-2 OS expresses a 3.8 kilobase mRNA which codes for the sequence of the PDGF-B chain.

#cross-references MUID:86164981
 #accession A25141

#molecule_type mRNA
 #residues 26-241 #label WEI
 #cross-references GB:X03702; NID:g35374; PID:g35375

REFERENCE A94271
 #authors Antoniadis, H.N.; Hunkapiller, M.W.
 #journal Science (1983) 220:963-965
 #title Human platelet-derived growth factor (PDGF): amino-terminal amino acid sequence.

#cross-references MUID:83197379
 #accession A94271

#molecule_type protein
 #residues 82-100, 'E', 102-104, 'C', 106, 'C', 108-110 #label ANT
 REFERENCE A93308

#authors Waterfield, M.D.; Scraze, G.T.; Whittle, N.; Stroobant, P.; Johnsson, A.; Wasteson, A.; Westermark, B.; Heldin, C.H.;

MPSRCH_PP protein - protein database search, using Smith-Waterman algorithm

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

(TM)

Title: >US-08-938-548A-6
Description: (1-130) from US08938548A.pep
Perfect Score: 954
Sequence: 1 MNLPTKVPWNAVTLILLL.....GRCPTATATAPRGGSRV 130
Scoring table: PAM 150
Gap 11
Searched: 63816 seqs, 5850866 residues
Post-processing: Minimum Match 04
Listing first 45 summaries
Database: a-issued
1:5_COMB 2:PCT9_COMB 3:backfiles
Statistics: Mean 27.785; Variance 127.171; scale 0.218
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Title: >US-08-938-548A-6									
Description: (1-130) from US08938548A.pep									
Perfect Score: 954									
Sequence: 1 MNLPSTKVPVAAVILLLL.....GRRCTATATALPRGSRV 130									
Scoring table: PAM 150									
Gap 11									
Searched: 63816 seqs, 5850866 residues									
Post-processing: Minimum Watch 0%									
Listing first 45 summaries									
Database: a-issued									
1:5_COMB 2:PCT9_COMB 3:backfiles									
Statistics: Mean 27.785; Variance 127.171; scale 0.218									
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.									
SUMMARIES									
Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.			
1	92	9.6	220	3	5175255-4 Patent No. 5175255.	4.65e+00			
2	89	9.3	241	3	5194596-15 Patent No. 5194596.	7.65e+00			
3	89	9.3	241	1	US-08-387- Sequence 4, Applicatio	7.65e+00			
4	89	9.3	241	2	PCT-US96-0 Sequence 9, Applicatio	7.65e+00			
5	88	9.2	254	1	US-08-236- Sequence 4, Applicatio	9.02e+00			
6	87	9.1	226	3	5498600-2 Patent No. 5498600.	1.06e+01			
7	87	9.1	241	3	5175255-2 Patent No. 5175255.	1.06e+01			
8	87	9.1	241	3	5175255-8 Patent No. 5175255.	1.06e+01			
9	87	9.1	282	1	US-08-445- Sequence 1, Applicati	1.06e+01			
10	85	8.9	241	1	US-08-469- Sequence 13, Applicati	1.47e+01			
11	85	8.9	613	2	PCT-US94-1 Sequence 2, Applicatio	1.47e+01			
12	84	8.8	238	1	US-08-240- Sequence 2, Applicatio	1.73e+01			
13	84	8.8	238	1	US-08-453- Sequence 2, Applicatio	1.73e+01			
14	84	8.8	1239	1	US-08-026- Sequence 3, Applicatio	1.73e+01			
15	83	8.7	235	2	PCT-US94-0 Sequence 6, Applicatio	2.04e+01			
16	83	8.7	235	1	US-08-243- Sequence 6, Applicatio	2.04e+01			
17	82	8.6	241	3	5219739-15 Patent No. 5219739.	2.39e+01			
18	80	8.4	492	1	US-07-794- Sequence 4, Applicatio	3.29e+01			
19	80	8.4	492	1	US-08-001- Sequence 2, Applicatio	3.29e+01			
20	80	8.4	1337	2	PCT-US95-0 Sequence 2, Applicatio	3.29e+01			
21	79	8.3	589	2	PCT-US91-0 Sequence 2, Applicatio	3.86e+01			
22	79	8.3	589	1	US-07-668- Sequence 2, Applicatio	3.86e+01			
23	78	8.2	1184	1	US-08-446- Sequence 20, Applicati	4.52e+01			

24	78	8.2	1184	1	US-08-446-	Sequence 20,	Applicati	4.52e+01
25	78	8.2	1187	1	US-08-097-	Sequence 13,	Applicati	4.52e+01
26	78	8.2	1187	1	US-08-357-	Sequence 8,	Applicati	4.52e+01
27	78	8.2	1187	2	PCT-US95-1	Sequence 8,	Applicatio	4.52e+01
28	77	8.1	831	1	US-08-073-	Sequence 5,	Applicatio	5.30e+01
29	77	8.1	831	1	US-08-483-	Sequence 5,	Applicatio	5.30e+01
30	77	8.1	831	1	US-08-234-	Sequence 5,	Applicatio	5.30e+01
31	77	8.1	2509	1	US-08-469-	Sequence 10,	Applicatio	5.30e+01
32	76	8.0	225	2	PCT-US94-0	Sequence 36,	Applicati	6.19e+01
33	76	8.0	219	3	5310729-4	Patent No.	5310729,	6.19e+01
34	76	8.0	231	1	US-08-243-	Sequence 2,	Applicatio	6.19e+01
35	76	8.0	231	1	US-08-220-	Sequence 7,	Applicatio	6.19e+01
36	76	8.0	231	2	PCT-US95-0	Sequence 6,	Applicatio	6.19e+01
37	76	8.0	231	2	PCT-US94-0	Sequence 2,	Applicatio	6.19e+01
38	76	8.0	263	2	PCT-US91-0	Sequence 2,	Applicatio	6.19e+01
39	76	8.0	488	1	US-07-794-	Sequence 2,	Applicatio	6.19e+01
40	76	8.0	488	1	US-08-001-	Sequence 2,	Applicatio	6.19e+01
41	76	8.0	529	2	PCT-US94-0	Sequence 2,	Applicatio	6.19e+01
42	76	8.0	4544	1	US-08-469-	Sequence 52,	Applicati	6.19e+01
43	75	7.9	171	1	US-08-313-	Sequence 4,	Applicatio	7.24e+01
44	75	7.9	447	1	US-07-937-	Sequence 29,	Applicati	7.24e+01
45	75	7.9	585	1	US-08-485-	Sequence 13,	Applicati	7.24e+01

ALIGNMENTS

RESULT 1

ID 5175255-4 STANDARD; PRT; 238 AA.

XX AC xxxxxx

XX 01-JAN-1900

XX DE Patent No. 5175255.

XX CC

CC Patent No. 5175255

CC APPLICANT: Thomason, Arlen R.; Nicholson, Margery

CC TITLE OF INVENTION: METHODS FOR PURIFICATION OF PLATELET-

CC DERIVED GROWTH FACTOR

CC NUMBER OF SEQUENCES: 9

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/06/25,344

CC FILING DATE: 23-MAR-1987

CC SEQ ID NO:4:

CC LENGTH: 220

CC SEQUENCE 238 AA; 26810 MW; 307752 CN;

Query Match 9.68; Score 92; DB 3; Length 220;

Best Local Similarity 46.2%; Pred. No. 4.65e+00;

Matches 18; Conservative 10; Mismatches 6; Indels 5; Gaps 5

Db 23 LQRLHGDPCGEAGAEIDLNNRSHSGELESIA-RGRR 60

QY 79 LQRLQAN-GNHA-AGI-LTMGR-RAGAELEPYPCGRR 113

|||||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

CC COMPUTER: Apple Power Macintosh
 CC OPERATING SYSTEM: Apple 7.5.3
 CC SOFTWARE: Microsoft Word, Version #6.0.1
 CC CURRENT APPLICATION DATA: US/08/236,918A
 CC APPLICATION NUMBER: 926,149
 CC FILING DATE: 05-AUG-1992
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 941,970
 CC FILING DATE: 15-DEC-1986
 CC APPLICATION NUMBER: 896,485
 CC FILING DATE: 3-AUG-1986
 CC APPLICATION NUMBER: 705,175
 CC FILING DATE: 25-FEB-1985
 CC APPLICATION NUMBER: 660,496
 CC FILING DATE: 12-OCT-1984
 CC SEQ ID NO:2:
 CC LENGTH: 226
 CC SEQUENCE 245 AA; 27563 MW; 324570 CN;

Query Match 9.2%; Score 88; DB 1; Length 254;
 Best Local Similarity 66.7%; Pred. No. 9.02e+00;
 Matches 12; Conservative 4; Mismatches 1; Indels 1; Gaps 1;
 Db 27 LPWLVAGLLLLLLLAAA 44
 QY 8 VFWAAVT-LLLLLLPPA 24

RESULT 6
 ID 5498600-2 STANDARD; PRT; 245 AA.
 XX XXXXXX
 AC XXXXXX
 DT 01-JAN-1900
 DE Patent No. 5498600.

Patent No. 5498600
 APPLICANT: MURRAY, MARK J.; KELLY, JAMES D.
 TITLE OF INVENTION: BIOLOGICALLY ACTIVE MOSAIC PROTEINS
 NUMBER OF SEQUENCES: 34
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/319,776
 FILING DATE: 07-OCT-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 926,149
 FILING DATE: 05-AUG-1992
 APPLICATION NUMBER: 379,239
 FILING DATE: 11-JUL-1989
 APPLICATION NUMBER: 941,970
 FILING DATE: 15-DEC-1986
 APPLICATION NUMBER: 896,485
 FILING DATE: 3-AUG-1986
 APPLICATION NUMBER: 705,175
 FILING DATE: 25-FEB-1985
 APPLICATION NUMBER: 660,496
 FILING DATE: 12-OCT-1984
 CC SEQ ID NO:2:
 CC LENGTH: 226
 CC SEQUENCE 245 AA; 27563 MW; 324570 CN;

Query Match 9.1%; Score 87; DB 3; Length 226;
 Best Local Similarity 43.6%; Pred. No. 1.06e+01;
 Matches 17; Conservative 11; Mismatches 6; Indels 5; Gaps 5;
 Db 29 LQRLQDGSKGEDGAEGLDNLNMTSRSHSGGELESIA-RGKR 66

QY 79 LQRLQAN-GNH-AAGI-LTMGR-RAGAELEPYPCGRR 113
 RESULT 7
 ID 5175255-2 STANDARD; PRT; 261 AA.
 XX XXXXXX
 AC XXXXXX
 DT 01-JAN-1900
 DE Patent No. 5175255.
 CC Patent No. 5175255
 CC APPLICANT: Thomason, Arlen R.; Nicholson, Margery
 CC TITLE OF INVENTION: METHODS FOR PURIFICATION OF PLATELET-
 CC DERIVED GROWTH FACTOR
 CC NUMBER OF SEQUENCES: 9
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/06/25,344
 CC FILING DATE: 23-MAR-1987
 CC SEQ ID NO:2:
 CC LENGTH: 241
 CC SEQUENCE 261 AA; 29370 MW; 369384 CN;

Query Match 9.1%; Score 87; DB 3; Length 241;
 Best Local Similarity 43.6%; Pred. No. 1.06e+01;
 Matches 17; Conservative 11; Mismatches 6; Indels 5; Gaps 5;
 Db 44 LQRLQDGSKGEDGAEGLDNLNMTSRSHSGGELESIA-RGKR 81
 QY 79 LQRLQAN-GNH-AAGI-LTMGR-RAGAELEPYPCGRR 113

RESULT 8
 ID 5175255-8 STANDARD; PRT; 261 AA.
 XX XXXXXX
 AC XXXXXX
 DT 01-JAN-1900
 DE Patent No. 5175255.

Patent No. 5175255
 APPLICANT: Thomason, Arlen R.; Nicholson, Margery
 TITLE OF INVENTION: METHODS FOR PURIFICATION OF PLATELET-
 CC DERIVED GROWTH FACTOR
 CC NUMBER OF SEQUENCES: 9
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/06/25,344
 CC FILING DATE: 23-MAR-1987
 CC SEQ ID NO:8:
 CC LENGTH: 241
 CC SEQUENCE 261 AA; 29326 MW; 367179 CN;
 Query Match 9.1%; Score 87; DB 3; Length 241;
 Best Local Similarity 43.6%; Pred. No. 1.06e+01;
 Matches 17; Conservative 11; Mismatches 6; Indels 5; Gaps 5;
 Db 44 LQRLQDGSKGEDGAEGLDNLNMTSRSHSGGELESIA-RGKR 81
 QY 79 LQRLQAN-GNH-AAGI-LTMGR-RAGAELEPYPCGRR 113

RESULT 9
 ID US-08-445-847A-1 STANDARD; PRT; 282 AA.
 XX XXXXXX
 AC XXXXXX
 DT
 DE
 XX Sequence 1, Application US/08445847A

CC LENGTH: 613 AMINO ACIDS
CC TYPE: AMINO ACID
CC STRANDEDNESS:
CC TOPOLOGY: LINEAR
CC MOLECULE TYPE: PROTEIN
SQ SEQUENCE 613 AA; 67034 MW; 1960405 CN;

Query Match 8.9%; Score 85; DB 2; Length 613;
Best Local Similarity 31.1%; Pred. No. 1.47e+01;
Matches 14; Conservative 13; Mismatches 18; Indels 0; Gaps 0;

Db 1 MRAPGALLARMLRLILLKLVSSASSALGVAPASRNCTLGESCA 45
QY 1 MNLPTKVPWAATLILLLLLPALLSLGVDAQPLDCCRQKTC 45

RESULT 12
ID US-08-240-124-2 STANDARD; PRT; 238 AA.

XX xxxxxx

Sequence 2, Application US/08240124

Sequence 2, Application US/08240124

Patent No. 5516658

GENERAL INFORMATION:

APPLICANT: BECKMANN, M. P.

APPLICANT: CERRETTI, DOUGLAS P.

TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE

TITLE OF INVENTION: RECEPTOR HEK

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: IMMUNEX CORPORATION

STREET: 51 UNIVERSITY STREET

CITY: SEATTLE

STATE: WASHINGTON

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple Macintosh

OPERATING SYSTEM: Apple System 7.1

SOFTWARE: Microsoft Word for Apple, Version 5.1a

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/240,124

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/161,132

FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/114,426

FILING DATE: 30-AUG-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/109,745

FILING DATE: 20-AUG-1993

ATTORNEY/AGENT INFORMATION:

NAME: SEESE, KATHRYN A.

REGISTRATION NUMBER: 32,172

REFERENCE/DOCKET NUMBER: 2814-C

TELEPHONE: (206) 587-0430

TELEFAX: (206) 233-0644

TELEX: 756822

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 238 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 238 AA; 26350 MW; 315668 CN;

Query Match 8.8%; Score 84; DB 1; Length 238;
Best Local Similarity 68.4%; Pred. No. 1.73e+01;
Matches 13; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

Db 2 AAAPLILLLLLVVPILLPL 20

QY 11 NAVTLLLLLLPP-ALLSL 28

RESULT 13
ID US-08-453-943-2 STANDARD; PRT; 238 AA.

XX xxxxxx

Sequence 2, Application US/08453943

Sequence 2, Application US/08453943

Patent No. 5738844

GENERAL INFORMATION:

APPLICANT: BECKMANN, M. P.

APPLICANT: CERRETTI, DOUGLAS P.

TITLE OF INVENTION: CYTOKINE THAT BINDS THE CELL SURFACE

TITLE OF INVENTION: RECEPTOR HEK

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: IMMUNEX CORPORATION

STREET: 51 UNIVERSITY STREET

CITY: SEATTLE

STATE: WASHINGTON

COUNTRY: USA

ZIP: 98101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Apple Macintosh

OPERATING SYSTEM: Apple System 7.1

SOFTWARE: Microsoft Word for Apple, Version 5.1a

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/453,943

FILING DATE: 30-MAY-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/240,124

FILING DATE: 09-MAY-1994

APPLICATION NUMBER: US 08/161,132

FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/114,426

FILING DATE: 30-AUG-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/109,745

FILING DATE: 20-AUG-1993

ATTORNEY/AGENT INFORMATION:

NAME: SEESE, KATHRYN A.

REGISTRATION NUMBER: 32,172

REFERENCE/DOCKET NUMBER: 2814-C

TELEPHONE: (206) 587-0430

TELEFAX: (206) 233-0644

TELEX: 756822

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 238 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 238 AA; 26350 MW; 315668 CN;

Query Match 8.8%; Score 84; DB 1; Length 238;
Best Local Similarity 68.4%; Pred. No. 1.73e+01;
Matches 13; Conservative 2; Mismatches 3; Indels 1; Gaps 1;

SQ SEQUENCE 235 AA; 26415 MW; 293990 CN;

Query Match 8.7%; Score 83; DB 2; Length 235;
Best Local Similarity 55.6%; Pred. No. 2.04e+01;
Matches 10; Conservative 6; Mismatches 1; Indels 1; Gaps 1;

Db 7 AWSPTTYLLLLLLSSGL 24
QY 9 PWAAYT-LLLLLLLPPAL 25

Search completed: Thu Jul 30 09:18:32 1998
Job time : 10 secs.

SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
[2]
RN SEQUENCE FROM N.A.
RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M., COULSON A.,
RA BRONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., FULTON L.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,
RL NATURE 368:32-38(0)
DR EMBL; 292970, E349623; -
DR PROSITE; PS00598; CHROMO_1; 2.
SQ SEQUENCE 1465 AA; 169391 MW; F959D52F CRC32;
Query Match 35.2%; Score 69; DB 3; Length 1465;
Best Local Similarity 69.2%; Pred. No. 2.48e+00;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
DB 700 QGRLLQQLKSSGK 712
|||||:|||||
QY 8 QGRLLQQLQASGN 20
RESULT 7
ID O31955 PRELIMINARY; PRT; 178 AA.
AC O31955;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE YONC PROTEIN.
GN YONC.
OS BACILLUS SUBTILIS.
OC PROKARYOTA; FIRMICUTES; ENDOSPORE-FORMING RODS AND COCCI; BACILLACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RA KUNST F., OGASAWARA N., MOSZER I., ALBERTINI A.M., ALLONI G.,
RA AZEVEDO V., BERTERO M.G., BESSIERES P., BOLOTIN A., BORCHERT S.,
RA BORRIS R., BOURSIER L., BRANS A., BRAUN M., BRIGHELL S.C., BRON S.,
RA BROUILLET S., BRUSCHI C.V., CALDWELL B., CAPUANO V., CARTER N.M.,
RA CHOI S.K., CODANI J.J., CONNERTON I.F., CUMMINGS N.J., DANIEL R.A.,
RA DENIZOT F., DEVINE K.M., DUSTERHOFT A., EHRLICH S.D., EMERSON P.T.,
RA ENTIAN K.D., ERINGTON J., FABRET C., FERRARI E., FOULGER D., FRITZ C.,
RA FUJITA M., FUJITA Y., FUMA S., GALIZZI A., GALLERON N., GHIM S.Y.,
RA GLASER P., GOFFEAU A., GOLIGHTLY E.J., GRANDI G., GUISEPPI G., GUY B.J.,
RA HAGA K., HAIBCH J., HARWOOD C.R., HENAUT A., HILBERT H., HOLSAPPEL S.,
RA HOSONO S., HULLO M.F., ITAYA M., JONES L., JORIS B., KARAMATA D.,
RA KASAHARA Y., KLAERR-BLANCHARD M., KLEIN C., KOBAYASHI Y., KOETTER P.,
RA KONGINGSTEIN G., KROGH S., KUMANO M., KURITA K., LAPIDUS A.,
RA LARDINOIS S., LAUBER J., LAZAREVIC V., LEE S.M., LEVINE A., LIU H.,
RA MASUDA S., MAUEL C., MEDIGUE C., MEDINA N., MELLADO R.P., MIZUNO M.,
RA MOWELL D., NAKAI S., NOBACK M., NOONE D., O'REILLY M., OGAWA K.,
RA OGWARA A., OUDEGA B., PARK S.H., PARRO V., POHL T.M., PORTELELL D.,
RA POWOOLLIK S., PRESCOTT A.M., PRESECAN E., PUJIC P., PURNELLE B.,
RA RAPPOULT G., REY M., REYNOLDS S., RIEGER M., RIVOLTA C., ROCHA E.,
RA ROCHE B., ROSE M., SADAIE Y., SATO T., SCANLAN E., SCHLEICH S.,
RA SCHROTER R., SCOFFONE F., SERIGUCHI J., SEKONKA A., SENOR S.J.,
RA SERROR P., SHIN B.S., SOLDO B., SOROKIN A., TACCONI E., TAKAGI T.,
RA TERPSTRA P., TOGNONI K., TAKEUCHI M., TAKAKOSHI A., TANAKA T.,
RA TAPAKASHI H., TAKEMARU K., TOSATO V., UCHIYAMA S., VANDENBOL M.,
RA VANNIER F., VASSAROTTI A., VIARI A., WAMBUUT R., WEDLER H.,
RA WEITZENEGER T., WINTERS P., WIPAT A., YAMAMOTO H., YAMANE K.,
RA YASUMOTO K., YATA K., YOSHIDA K., YOSHIKAWA H.F., ZUMSTEIN E.,
RL NATURE 390:249-256(1997).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-168;
RA KUNST F., OGASAWARA N., YOSHIKAWA H., DANCHIN A.;

RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; 299115; E1183561; -
SQ SEQUENCE 178 AA; 19609 MW; CA7B8ED0 CRC32;
Query Match 34.7%; Score 68; DB 9; Length 178;
Best Local Similarity 45.0%; Pred. No. 3.53e+00;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
DB 1 MATRLQKALTEVGNHTGNL 20
: |||: | |||: |
QY 7 LQRLQRLQASGNHAAGIL 26
RESULT 8
ID Q60811 PRELIMINARY; PRT; 580 AA.
AC Q60811;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE RETINOID X RECEPTOR INTERACTING PROTEIN 110 (FRAGMENT).
GN RXRIP110 OR RIP110.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE; 95280959.
RA SEOL W., CHOI H.S., MOORE D.D.;
RL MOL. ENDOCRINOL. 9:72-85(1995).
DR EMBL; U22015; G709961; -
DR MGD; MGI:103185; RXRIP110.
FT NON_TER 1
SQ SEQUENCE 580 AA; 64852 MW; 28DA7406 CRC32;
Query Match 34.7%; Score 68; DB 10; Length 580;
Best Local Similarity 44.4%; Pred. No. 3.53e+00;
Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
DB 485 QGRLLSLLEQSEHRTTG 502
|||||: |||: | :|||: |
QY 8 QGRLLQQLQASGNHAAGI 25
RESULT 9
ID P95644 PRELIMINARY; PRT; 462 AA.
AC P95644;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE RNA POLYMERASE SIGMA FACTOR.
GN RPOD.
OS RENIBACTERIUM SALMONINARUM.
OC EUBACTERIA; FIRMICUTES; ACTINOMYCETES; RENIBACTERIUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC33209;
RA BARRIL A., CONCHA M.I., FIGUEROA J., LEON G.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC THEN IS RELEASED.
DR EMBL; Y10835; E294130; -
DR PROSITE; PS00715; SIGMA70_1; 1.
DR PROSITE; PS00716; SIGMA70_2; 1.
KW TRANSCRIPTION REGULATION; SIGMA FACTOR; DNA-DIRECTED RNA POLYMERASE;
KW DNA-BINDING.
SQ SEQUENCE 462 AA; 50504 MW; E54BBF5 CRC32;
Query Match 34.2%; Score 67; DB 9; Length 462;
Best Local Similarity 31.8%; Pred. No. 5.01e+00;
Matches 7; Conservative 11; Mismatches 4; Indels 0; Gaps 0;

DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DE 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DS POLYPROTEIN.
OS BIRNAVIRUS SP.
OC VIRUSES; DSRNA VIRUSES; BIRNAVIRIDAE; BIRNAVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Y-6;
RA SUZUKI S., KIMURA M., KUSUDA R.;
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Y-6;
RA SUZUKI S., KIMURA M., KUSUDA R.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AB001329; D1019990; -;
DR EMBL; AB006783; D1022914; -;
KW POLYPROTEIN.
SQ SEQUENCE 972 AA; 106767 MW; 9905CA0D CRC32;

Query Match 34.28; Score 67; DB 11; Length 972;
Best Local Similarity 32.08; Pred. No. 5.01e+00;
Matches 8; Conservative 10; Mismatches 7; Indels 0; Gaps 0;

Db 731 RTTASGMDEELQKLLHATMARAKEV 755
QY 1 RSGPPLGQRLQLQASGNHAGI 25
| : : : : | : : : : | : : : : | : : : : |

RESULT 15
ID 033958 PRELIMINARY; PRT; 1841 AA.
AC 033958;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DE 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DS TYLACTONE SYNTHASE MODULE 7.
GN TYLG
OS STREPTOMYCES FRADIAE.
OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; STREPTOMYCETACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RA DEHOFF B.S., SUTTON K.L., ROSTECK P.R. JR.;
RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U78289; G2317864; -;
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
KW TRANSFERASE.
SQ SEQUENCE 1841 AA; 191190 MW; 2FF1A058 CRC32;

Query Match 34.28; Score 67; DB 9; Length 1841;
Best Local Similarity 33.38; Pred. No. 5.01e+00;
Matches 8; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

Db 1513 RDTPAALAAHLAELLATARDHGP 1536
QY 1 RSGPPLGQRLQLQASGNHAG 24
| : : : : | : : : : | : : : : | : : : : |

Search completed: Thu Jul 30 09:12:50 1998
Job time : 23 secs.

W P E R L H

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm

on: Thu Jul 30 09:14:41 1998; MasPar time 5.30 Seconds
376.024 Million cell updates/sec
ular output not generated.

Title: >US-08-938-548A-6
Description: (1-130) from US08938548A.pep
Perfect Score: 954
Sequence: 1 MNPSTKVPWAAVTLILL.....GRRCTATATAPRGSSRV 130

Scoring table: PAM 150
Gap 11

Searched: 124785 seqs, 15338987 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq31-2
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27

Statistics: Mean 29.913; Variance 129.900; scale 0.230

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	96	10.1	566	2 R11148	Fibulin A.	5.91e+00
2	96	10.1	601	2 R11149	Fibulin B.	5.91e+00
3	96	10.1	683	2 R11150	Fibulin C.	5.91e+00
4	91	9.5	440	2 P70134	Natural recombinant h	1.39e+01
5	91	9.5	440	25 W24789	Human lecithin-choles	1.39e+01
6	89	9.3	190	9 R50012	Truncated Platelet de	1.94e+01
7	89	9.3	216	12 R68617	Human PDGF-B precurs	1.94e+01
8	89	9.3	220	8 R40965	rPDGF B.	1.94e+01
9	89	9.3	241	9 R50009	Platelet-derived Grow	1.94e+01
10	89	9.3	241	12 R63469	Recombinant platelet	1.94e+01
11	89	9.3	241	9 R50002	Platelet-derived Grow	1.94e+01
12	89	9.3	241	8 R40964	PDGF Bc-sis.	1.94e+01
13	89	9.3	241	1 R60536	Recombinant platelet	1.94e+01
14	88	9.2	252	4 R21708	HSV-1 (CVG-2) ICP34.5	2.30e+01
15	88	9.2	254	12 R64190	Human 4-1BB-L polypep	2.30e+01
16	88	9.2	254	25 W26657	Human 4-1BB ligand.	2.30e+01
17	87	9.1	111	7 R35515	Tryptophan aporepress	2.72e+01
18	87	9.1	190	12 R60615	Human PDGF-B 109 subu	2.72e+01
19	87	9.1	201	4 R21443	Sequence encoded by p	2.72e+01

20	87	9.1	226	2 R22673	v-sis protein p38sis.	2.72e+01
21	87	9.1	241	1 R80597	Cv-sis gene encoded p	2.72e+01
22	87	9.1	241	12 R63472	Recombinant platelet	2.72e+01
23	87	9.1	241	8 R40967	Cv-sis gene product.	2.72e+01
24	87	9.1	271	1 R80595	Recombinant platelet	2.72e+01
25	87	9.1	271	12 R63468	Recombinant platelet	2.72e+01
26	87	9.1	271	8 R40963	PDGF Bv-sis.	2.72e+01
27	87	9.1	282	12 R60616	Human PDGF-B 119 link	2.72e+01
28	87	9.1	599	4 R21690	Prostaglandin endoper	2.72e+01
29	85	8.9	226	12 R63470	Recombinant platelet	3.79e+01
30	85	8.9	226	3 P81029	Sequence encoded by t	3.79e+01
31	85	8.9	226	3 P60215	Sequence encoded by t	3.79e+01
32	85	8.9	363	23 W12414	Porcine complement in	3.79e+01
33	85	8.9	613	18 R98261	Human endothelin-bomb	3.79e+01
34	84	8.8	35	1 R94256	Truncated E protein f	4.47e+01
35	84	8.8	234	16 R82605	Eph transmembrane tyr	4.47e+01
36	84	8.8	238	13 R71481	Human hek-L protein.	4.47e+01
37	84	8.8	1239	9 R45945	Glutamic acid recepto	4.47e+01
38	83	8.7	235	13 R66175	Human S86/SL09 Flt3 l	5.26e+01
39	83	8.7	235	12 R67541	Human flt-3 ligand.	5.26e+01
40	83	8.7	551	13 R77858	S. clavuligerus ORF1	5.26e+01
41	82	8.6	35	1 P91354	Amino acid sequence o	6.20e+01
42	82	8.6	241	7 R38919	Human Platelet Derive	6.20e+01
43	81	8.5	152	27 W29307	Wild-type avidin prot	7.29e+01
44	81	8.5	438	27 W37355	Wzy (Wfc) protein inv	7.29e+01
45	80	8.4	1337	14 R85203	huDEP-1.	8.58e+01

ALIGNMENTS

RESULT 1
ID R11148 standard; Protein; 566 AA.
AC R11148;
DT 21-MAY-1991 (first entry)
DE Fibulin A.
KW Beta-1 integrin; adhesion; receptor; fibronectin.
OS Homo sapiens.
FH peptide
FT 1..29 Location/Qualifiers
FT protein 30..566 /label= signal sequence
FT modified_site 98 /label= fibulin A
FT modified_site 535 /label= N-linked glycosylation
FT modified_site 339 /label= N-linked glycosylation
FT region 36..144 /label= N-linked glycosylation
FT region 36..69 /label= type I motif
FT region 112..144 /label= repeat unit 1
FT region 144..179 /label= repeat unit 2
FT region 180..566 /label= Glu/Asp-rich region
FT region 180..214 /label= type II motif
FT region 215..219 /label= repeat unit 1
FT region 220..260 /label= consensus pentapeptide
FT region 261..265 /label= repeat unit 2
FT region 266..306 /label= consensus pentapeptide
FT region 307..311 /label= repeat unit 3
FT region 312..354 /label= consensus pentapeptide
FT region 355..359 /label= repeat unit 4

MPSRH
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:11:23 1998; MasPar time 4.07 Seconds
251.156 Million cell updates/sec

ular output not generated.

Title: >US-08-938-548A-4
Description: (1-28) from US08938548A.pap
Perfect Score: 196
Sequence: 1 RSGPPGLQRLQLLQASGHNHAGILTM 28

Scoring table: PAM 150
Gap 15

Searched: 120441 seqs, 36531193 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: p1r56
1:p1r1 2:p1r2 3:p1r3 4:p1r4 5:nrl3d

Statistics: Mean 30.060; Variance 54.689; scale 0.550

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	69	35.2	439	2	S61858 hrpE protein - Pseudo	3.92e+00
2	69	35.2	736	2	S47044 MALDP protein - mouse	3.92e+00
3	68	34.7	178	2	E69913 hypothetical protein	5.47e+00
4	67	34.2	442	2	JH0496 transcription initiat	7.60e+00
5	67	34.2	442	2	S11712 transcription initiat	7.60e+00
6	67	34.2	510	2	S41307 transcription initiat	7.60e+00
7	67	34.2	525	2	JN0443 transcription initiat	7.60e+00
8	67	34.2	528	2	JN0445 transcription initiat	7.60e+00
9	67	34.2	1477	2	S64616 YOR1 protein - yeast	7.60e+00
10	67	34.2	2115	2	S23647 NUMA protein - human	7.60e+00
11	66	33.7	451	2	S36593 E2 protein - human pa	1.05e+01
12	65	33.2	178	2	S26481 calcium-binding prote	1.45e+01
13	65	33.2	1692	2	A33988 adenylate cyclase (EC	1.45e+01
14	64	32.7	263	2	S51155 hypothetical protein	2.00e+01
15	64	32.7	325	1	QOBE38 BGRF1 protein - human	2.00e+01
16	64	32.7	394	2	A42115 ribosomal protein S2,	2.00e+01
17	64	32.7	400	2	S76929 hypothetical protein	2.00e+01
18	64	32.7	535	2	A46101 protein-tyrosine-phos	2.00e+01
19	64	32.7	548	2	B46101 protein-tyrosine-phos	2.00e+01
20	64	32.7	556	2	B64939 hypothetical protein	2.00e+01
21	64	32.7	879	2	H64888 hypothetical protein	2.00e+01
22	63	32.1	248	1	LNHUP5 pulmonary surfactant	2.74e+01
23	63	32.1	248	1	LNHUP1 pulmonary surfactant	2.74e+01

24	63	32.1	248	1	LNHUP6 pulmonary surfactant	2.74e+01
25	63	32.1	248	2	I51921 pulmonary surfactant-	2.74e+01
26	63	32.1	256	1	O3YCRQ hypothetical protein	2.74e+01
27	63	32.1	324	1	LURT3 annexin III - rat	2.74e+01
28	63	32.1	331	2	S76765 hypothetical protein	2.74e+01
29	63	32.1	429	2	S09852 hypothetical protein	2.74e+01
30	63	32.1	878	2	A41055 ecdysone receptor - f	2.74e+01
31	63	32.1	972	1	GNXSIV genome polypotein -	2.74e+01
32	63	32.1	1157	2	A55152 PAS1 protein - yeast	2.74e+01
33	62	31.6	36	2	B49139 allergen TBA-1 - Toxo	3.75e+01
34	62	31.6	187	2	S48361 transcription factor	3.75e+01
35	62	31.6	253	2	S49183 hypothetical protein	3.75e+01
36	62	31.6	281	2	B64899 hypothetical protein	3.75e+01
37	62	31.6	319	1	SL9248 RNA-directed DNA poly	3.75e+01
38	62	31.6	345	2	F69200 conserved hypothetical	3.75e+01
39	62	31.6	511	2	JC2294 carboxypeptidase (PC	3.75e+01
40	62	31.6	624	1	RDYCS7 sulfite reductase (fe	3.75e+01
41	62	31.6	972	1	GNXSN1 genome polypotein -	3.75e+01
42	62	31.6	1176	2	A49848 nitrite reductase (NA	3.75e+01
43	62	31.6	2504	2	A57788 fatty-acid synthase (3.75e+01
44	62	31.6	2504	2	B57788 fatty-acid synthase (3.75e+01
45	62	31.6	2509	2	G01880 fatty-acid synthase (3.75e+01

ALIGNMENTS

RESULT 1

ENTRY S61858 #type complete

TITLE hrpE protein - Pseudomonas solanacearum

ORGANISM #formal_name Pseudomonas solanacearum

DATE 20-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 10-Sep-1997

ACCESSIONS S61858

REFERENCE S61846

#authors van Gijsegem, F.; Gough, C.; Zischek, C.; Niqueux, E.; Ariat, M.; Genin, S.; Barberis, P.; German, S.; Castello, P.; Boucher, C.

#journal Mol. Microbiol. (1995) 15:1095-1114

#title The hrp gene locus of Pseudomonas solanacearum, which controls the production of a type III secretion system, encodes eight proteins related to components of the bacterial flagellar biogenesis complex.

#accession S61858

##status preliminary; nucleic acid sequence not shown

##molecule_type DNA

##residues 1-439 #label VAN

##cross-references EMBL:Z14056; NID:G550397; PID:G550413

GENETICS

#gene hrpE

CLASSIFICATION #superfamily H+-transporting ATP synthase alpha chain homology

FEATURE 191-359 #domain H+-transporting ATP synthase alpha chain homology #label ATP

SUMMARY #length 439 #molecular-weight 47533 #checksum 9718

Query Match 35.2%; Score 69; DB 2; Length 439;
Best Local Similarity 44.4%; Pred. No. 3.92e+00;
Matches 12; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

Db 281 RGPFPVFAELPRLERAGMSAGSIT 307
| : | | : : | | | : | | | : |
QY 1 RSGPPGLQRLQLLQASGHNHAGILT 27

RESULT 2

ENTRY S47044 #type complete

TITLE MALDP protein - mouse

ORGANISM #formal_name Mus musculus

DATE 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 10-Sep-1997

ACCESSIONS S47044

REFERENCE S47044

```

#cross-references EMBL:X52983; NID:g48744; PID:g48745
#experimental_source strain A3(2)
GENETICS
#gene hrdB
CLASSIFICATION #superfamily transcription initiation factor sigma katF;
transcription initiation factor sigma katF homology
KEYWORDS DNA binding; sigma factor; transcription initiation
FEATURE
211-437 #domain transcription initiation factor sigma katF
homology #label KTF
SUMMARY #length 442 #molecular-weight 48413 #checksum 4794
Query Match 34.2%; Score 67; DB 2; Length 442;
Best Local Similarity 31.8%; Pred. No. 7.60e+00;
Matches 7; Conservative 11; Mismatches 4; Indels 0; Gaps 0;
Db 370 LQQLHSVLDLTLSEAGVYSM 391
|| :|: :|: :|: :|: :|:
QY 7 LQGLRQLLQASGNHAAIGLTM 28

ULT 6
ENTRY #S41307 #type complete
TITLE transcription initiation factor sigma - Streptomyces griseus
ORGANISM #formal_name Streptomyces griseus
DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
08-Sep-1997
ACCESSIONS S41307
REFERENCE #authors Marcos, A.T.; Diez, B.; Gutierrez, S.; Fernandez, F.J.;
Velasco, J.; Martin, J.F.
#submission Submitted to the EMBL Data Library, December 1993
#description Organization and expression of the hrdB-sprc gene cluster of
Streptomyces griseus encoding a sigma factor protein and a
serine protease. Role on growth and sporulation of
Streptomyces.
#accession S41307
#status preliminary
#molecule_type DNA
#residues 1-510 #label MAR
#cross-references EMBL:X75952; NID:g440164; PID:g581664
GENETICS
#start_codon GTG
CLASSIFICATION #superfamily transcription initiation factor sigma katF;
transcription initiation factor sigma katF homology
KEYWORDS DNA binding; sigma factor; transcription initiation
FEATURE
279-505 #domain transcription initiation factor sigma katF
homology #label KTF
SUMMARY #length 510 #molecular-weight 55795 #checksum 1415
Query Match 34.2%; Score 67; DB 2; Length 510;
Best Local Similarity 31.8%; Pred. No. 7.60e+00;
Matches 7; Conservative 11; Mismatches 4; Indels 0; Gaps 0;
Db 438 LQQLHSVLDLTLSEAGVYSM 459
|| :|: :|: :|: :|: :|:
QY 7 LQGLRQLLQASGNHAAIGLTM 28

ULT 7
ENTRY #JN0443 #type complete
TITLE transcription initiation factor sigma homolog hrdB -
Streptomyces aureofaciens
ORGANISM #formal_name Streptomyces aureofaciens
DATE 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
08-Sep-1997
ACCESSIONS JN0443
REFERENCE #authors Kormanec, J.; Parkasovsky, M.; Poutuckova, L.
#journal Gene (1992) 122:63-70
#title Four genes in Streptomyces aureofaciens containing a domain
characteristic of principal sigma factors.
#accession JN0443
#molecule_type DNA
#residues 1-528 #label KOR
#cross-references GB:M90412; NID:g153308; PID:g153309
GENETICS
#gene hrdE
#start_codon GTG
CLASSIFICATION #superfamily transcription initiation factor sigma katF;
transcription initiation factor sigma katF homology
KEYWORDS DNA binding; sigma factor; transcription initiation
FEATURE
297-523 #domain transcription initiation factor sigma katF
homology #label KTF
SUMMARY #length 528 #molecular-weight 57598 #checksum 7360
Query Match 34.2%; Score 67; DB 2; Length 528;
Best Local Similarity 36.4%; Pred. No. 7.60e+00;
Matches 8; Conservative 9; Mismatches 5; Indels 0; Gaps 0;
Db 456 LQQLHSVLDLTLSEAGVYSM 477
|| :|: :|: :|: :|: :|:
QY 7 LQGLRQLLQASGNHAAIGLTM 28

ULT 9
ENTRY #S64616 #type complete
TITLE YOR1 protein - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES protein G9537; protein YGR281w
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 17-May-1996 #sequence_revision 17-May-1996 #text_change
20-Feb-1998
ACCESSIONS S64616
REFERENCE #authors Voet, M.; Voickaert, G.
#submission submitted to the Protein Sequence Database, May 1996
#accession S64616
#molecule_type DNA

```

```
#title      Adenylate cyclases in yeast: a comparison of the genes from
#cross-references MUID:89345533
#accession  A35539
#status     preliminary
#molecule_type DNA
#residues   1-1692 #label YAM
#cross-references GB:M24942; NID:g173378; PID:g173379
#note       the authors translated the codon TGC for residue 626 as
            Ser, and GCC for residue 1243 as Gly
CLASSIFICATION #superfamily leucine-rich alpha-2-glycoprotein repeat
                homology; yeast adenylate cyclase catalytic domain homology
                phosphorus-oxygen lyase
KEYWORDS      #domain yeast adenylate cyclase catalytic domain
FEATURE       1328-1413
SUMMARY       homology #label YACC
                #length 1692 #molecular-weight 190332 #checksum 2609
Query Match   33.2%; Score 65; DB 2; Length 1692;
Best Local Similarity 57.1%; Pred. No. 1.45e+01;
Matches       8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
DB 1539 PVQLQRLERLIKS 1552
      1 11111111:
QY 4 PPLQGLRLQLLQA 17

RESULT 14
ENTRY   S51155 #type complete
TITLE   hypothetical protein 263 - Brassica tournefortii
ORGANISM mitochondrion
DATE     #formal_name mitochondrion Brassica tournefortii
        19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change
        10-Sep-1997
ACCESSIONS S51155
REFERENCE   S51154
#authors   Landgren, M.; Zetterstrand, M.; Sundberg, E.; Glimelius, K.
#submission submitted to the EMBL Data Library, January 1995
#description Mitochondrial analysis revealed transcription of an ORF 3 of
            the atp6-gene and translation of a 32 kD protein in
            alloplasmic male sterile Brassica lines containing the B.
            tournefortii cytoplasm.
#accession S51155
#molecule_type DNA
#residues  1-263 #label LAN
#cross-references EMBL:X83692; NID:g633155; PID:g633157
GENETICS
#genome   mitochondrion
#WORDS     mitochondrion
#MARY      #length 263 #molecular-weight 29082 #checksum 9529
Query Match   32.7%; Score 64; DB 2; Length 263;
Best Local Similarity 45.0%; Pred. No. 2.00e+01;
Matches       9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
DB 1 MNRLQWLPLLGSSVAGFL 20
      : 11111111:
QY 7 LQGLRLQLLQASGNHAAGIL 26

RESULT 15
ENTRY   Q08E38 #type complete
TITLE   BGRF1 protein - human herpesvirus 4 (strain B95-8)
ORGANISM #formal_name human herpesvirus 4, Epstein-Barr virus
DATE     25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change
        05-Sep-1997
ACCESSIONS A43044; A03782; S33034
REFERENCE   A93065
#authors   Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
#journal   Mol. Biol. Med. (1983) 1:21-45
#title     Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8
            Epstein-Barr virus.
#cross-references MUID:85035713
```

```
#accession  A43044
#molecule_type DNA
#residues   1-325 #label BAN
#cross-references EMBL:V01555; NID:g59074; PID:g1334893
REFERENCE   A03794
#authors   Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.;
            Farrell, P.J.; Gibson, T.J.; Hatfull, G.; Hudson, G.S.;
            Satchwell, S.C.; Sequin, C.; Tuffnell, P.S.; Barrell, B.G.
#journal   Nature (1984) 310:207-211
#title     DNA sequence and expression of the B95-8 Epstein-Barr virus
            genome.
#cross-references MUID:84270667
#contents  annotation: protein coding region
CLASSIFICATION #superfamily herpesvirus 38K protein
SUMMARY       #length 325 #molecular-weight 36462 #checksum 3547
Query Match   32.7%; Score 64; DB 1; Length 325;
Best Local Similarity 39.1%; Pred. No. 2.00e+01;
Matches       9; Conservative 4; Mismatches 10; Indels 0; Gaps 0;
DB 104 PELRDTLQRLPPPPNLEDEALT 126
      1 11111111:
QY 5 PGLQGLRLQLLQASGNHAAGILT 27
```

Search completed: Thu Jul 30 09:11:43 1998
Job time : 20 secs.

FT CARBOHYD 218 218 POTENTIAL.
FT CARBOHYD 295 295 POTENTIAL.
FT CARBOHYD 661 661 POTENTIAL.
FT CARBOHYD 759 759 POTENTIAL.
FT CARBOHYD 799 799 POTENTIAL.
FT CARBOHYD 1345 1345 POTENTIAL.
FT CARBOHYD 1366 1366 POTENTIAL.
SQ SEQUENCE 1477 AA; 166727 MW; 79B302B8 CRC32;

Query Match 34.2%; Score 67; DB 1; Length 1477;
Best Local Similarity 40.0%; Pred. No. 2.34e+00;
Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Db 499 GRLOSLEAPDDPNQMIEM 518
||| |||: : : : :
QY 9 GRLORLQASGNNHAAGILTM 28

RESULT 6
ID VE2.HPV09 STANDARD; PRT; 461 AA.
P36780;
01-JUN-1994 (REL. 29, CREATED)
01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE REGULATORY PROTEIN E2.
GN E2.
OS HUMAN PAPILLOMAVIRUS TYPE 9.
OC VIRIDAE; DS-DNA NONENVELOPED VIRUSES; PAPAPOVIRIDAE; PAPILLOMAVIRUSES.
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 94265501.
RA DELIUS H., HOFMANN B.;
RL CURR. TOP. MICROBIOL. IMMUNOL. 186:13-31(1994).
CC -!- FUNCTION: E2 IS A TRANSCRIPTIONAL TRANSACTIVATOR CAPABLE OF
ACTIVATING A CONDITIONAL ENHANCER IN THE VIRAL LONG CONTROL
CC REGION (LCR). E2 BINDS TO THE 5'-ACCGNCGGT-3' PALINDROMIC
CC SEQUENCE.
CC -!- SUBUNIT: BINDS DNA AS A DIMER.
DR EMBL; X74464; G397073; -.
DR PIR; S36593; S36593.
DR HSP; P11299; 2BOP.
KW EARLY PROTEIN; TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING;
KW TRANS-ACTING FACTOR.
SQ SEQUENCE 461 AA; 52141 MW; DA2B4125 CRC32;

Query Match 33.7%; Score 66; DB 1; Length 461;
Best Local Similarity 38.5%; Pred. No. 3.39e+00;
Matches 10; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

358 RSVGAGHGHRLARLAEAKDPPLMLL 383
|| : ||| ||| : : : :
QY 1 RSGPPGLOGRLQRLQASGNNHAAGIL 26

RESULT 7
ID CYAA.SCHPO STANDARD; PRT; 1692 AA.
AC P14605;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
DE ADENYLATE CYCLASE (EC 4.6.1.1) (ATP PYROPHOSPHATE-LYASE) (ADENYLATE
DE CYCLASE).
GN CYRL.
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCUMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90046723.
RA YOUNG D., RIGGS M., FIELD J., VOUTEK A., BROEK D., WIGLER M.;
RL PROC. NATL. ACAD. SCI. U.S.A. 86:7989-7993(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 89345533.

RA YAMAWAKI-KATAOKA Y., TAMAOKI T., CHOE H.-R., TANAKA H., KATAOKA T.;
RL PROC. NATL. ACAD. SCI. U.S.A. 86:5693-5697(1989).
CC -!- FUNCTION: PLAYS ESSENTIAL ROLES IN REGULATION OF CELLULAR
METABOLISM BY CATALYSING THE SYNTHESIS OF A SECOND MESSENGER,
CC CAMP.
CC -!- CATALYTIC ACTIVITY: ATP -> 3',5'-CYCLIC AMP + PYROPHOSPHATE.
CC -!- IN CONTRAST TO YEAST CYCLASE, S. POMBE CYCLASE IS NOT LIKELY TO BE
CC REGULATED BY RAS PROTEINS.
CC -!- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN
CC MANY PROTEINS. NUMBER IN THIS PROTEIN: 22.
CC -!- SIMILARITY: BELONGS TO CLASS-3 OF ADENYLATE CYCLASES.
CC -!- SIMILARITY: CONTAINS A PP2C-LIKE DOMAIN.
DR EMBL; M26699; G173339; -.
DR EMBL; M24942; G173379; -.
DR PIR; A33988; A33988.
DR PIR; A33539; A33539.
KW LYASE; REPEAT; LEUCINE-REPEAT; CAMP SYNTHESIS; MAGNESIUM.
FT DOMAIN 326 999 LEUCINE-RICH REPEATS.
FT REPEAT 326 349 LRR 1.
FT REPEAT 350 396 LRR 2.
FT REPEAT 397 420 LRR 3.
FT REPEAT 421 444 LRR 4.
FT REPEAT 445 466 LRR 5.
FT REPEAT 467 492 LRR 6.
FT REPEAT 493 517 LRR 7.
FT REPEAT 518 540 LRR 8.
FT REPEAT 541 563 LRR 9.
FT REPEAT 564 587 LRR 10.
FT REPEAT 588 616 LRR 11.
FT REPEAT 617 674 LRR 12.
FT REPEAT 675 698 LRR 13.
FT REPEAT 699 722 LRR 14.
FT REPEAT 723 744 LRR 15.
FT REPEAT 745 798 LRR 16.
FT REPEAT 799 822 LRR 17.
FT REPEAT 823 852 LRR 18.
FT REPEAT 853 892 LRR 19.
FT REPEAT 893 922 LRR 20.
FT REPEAT 923 950 LRR 21.
FT REPEAT 951 999 LRR 22.
FT DOMAIN 1008 1276 PP2C-LIKE.
FT DOMAIN 1277 1692 CATALYTIC.
SQ SEQUENCE 1692 AA; 190333 MW; 20E7D7D8 CRC32;

Query Match 33.2%; Score 65; DB 1; Length 1692;
Best Local Similarity 57.1%; Pred. No. 4.88e+00;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 1539 PVOLQGLRLERLIKS 1552
| |||||: : : : :
QY 4 PPGLOGRLQRLQQA 17

RESULT 8
ID RT04.YEAST STANDARD; PRT; 394 AA.
AC F32902;
DT 01-OCT-1993 (REL. 27, CREATED)
DT 01-OCT-1993 (REL. 27, LAST SEQUENCE UPDATE)
DT 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE MITOCHONDRIAL 40S RIBOSOMAL PROTEIN MRP4.
GN MRP4 OR YH1004W.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCUMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92184810.
RA DAVIS S.C., TZAGOLOFF A., ELLIS S.R.;
RL J. BIOL. CHEM. 267:5508-5514(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / AB972;
RX MEDLINE; 94378003.
RA JOHNSTON M., ANDREWS S., BRINKMAN R., COOPER J., DING H., DOVER J.,

FT CONFLICT 36 36 I -> L (IN REF. 3).
SQ SEQUENCE 879 AA; 96834 MW; 8BFD7CF3 CRC32;

Query Match 32.7%; Score 64; DB 1; Length 879;
Best Local Similarity 45.8%; Pred. No. 7.01e+00;
Matches 11; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

Db 396 GVDGRQLQALQAHNELGDFVLHM 419
QY 6 GLQGRQLRLQASGNHAG-ILTM 28

RESULT 12

ID IL11 MOUSE STANDARD; PRT; 199 AA.

AC P47873;

DT 01-FEB-1996 (REL. 33, CREATED)

DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE INTERLEUKIN-11 PRECURSOR (IL-11).

GN IL11.

MUS MUSCULUS (MOUSE).

EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

[1]

SEQUENCE FROM N.A.

RX MEDLINE; 97070356.

RA MORRIS J.C., FINERTY H., BENNETT F., TURNER K.J., WOOD C.R.;

RL EXP. HEMATOL. 24:1369-1376(1996).

CC -1- FUNCTION: THIS PROTEIN STIMULATES PLASMACYTOMA PROLIFERATION,

CC T-CELL-DEPENDENT DEVELOPMENT OF IMMUNOGLOBULIN-PRODUCING B

CC CELLS AND SYNERGIZES WITH IL-3 IN SUPPORTING MURINE

CC MEGAKARYOCYTE COLONY FORMATION (BY SIMILARITY).

DR EMBL; U03421; G415654; -.

DR MGD; MGI:107613; IL11.

KW CYTOKINE; GROWTH FACTOR; SIGNAL.

FT SIGNAL 1 21

FT CHAIN 22 199 INTERLEUKIN-11.

SQ SEQUENCE 199 AA; 21522 MW; 1CB30772 CRC32;

Query Match

Best Local Similarity 32.1%; Score 63; DB 1; Length 199;

Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 128 ALQARLERILR 138

QY 6 GLQGRQLRLQ 16

ULT 13

PSPA_HUMAN STANDARD; PRT; 248 AA.

P07714;

DT 01-APR-1988 (REL. 07, CREATED)

DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)

DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE PULMONARY SURFACTANT-ASSOCIATED PROTEIN A PRECURSOR (SP-A) (PSP-A)

DE (PSP) (ALVEOLAR PROTEINOSIS PROTEIN).

GN SFTPA OR SFTPI OR PSPA.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

OC EUTHERIA; PRIMATES.

[1]

SEQUENCE FROM N.A.

RX MEDLINE; 86250832.

RA FLORES J., STEINBRINK R., JACOBS K., PHELPS D., KRIZ R., RECHNY M.,

RA SULTZMAN L., JONES S., TAEUSCH H.W., FRANK H.A., FRITSCH E.F.;

RL J. BIOL. CHEM. 261:9029-9033(1986).

[2]

SEQUENCE FROM N.A.

RX MEDLINE; 86014366.

RA WHITE R.T., DAMM D., MILLER J., SPRATT K., SCHILLING J., HAWGOOD S.,

RA BENSON B., CORDELL B.;

RL NATURE 317:361-363(1985).

[3]

RP SEQUENCE FROM N.A.
RX MEDLINE; 92198680.
RA KATVAL S.L., SINGH G., LOCKER J.L.;

RL AM. J. RESPIR. CELL MOL. BIOL. 6:446-452(1992).

CC -1- FUNCTION: IN PRESENCE OF CALCIUM IONS, PSAP BINDS TO SURFACTANT

CC PHOSPHOLIPIDS AND CONTRIBUTES TO LOWER THE SURFACE TENSION AT THE

CC AIR-LIQUID INTERFACE IN THE ALVEOLI OF THE MAMMALIAN LUNG AND IS

CC ESSENTIAL FOR NORMAL RESPIRATION.

CC -1- SUBUNIT: OLIGOMERIC COMPLEX OF 6 SET OF HOMOTRIMERS.

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.

CC -1- PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10% PROTEIN. THERE

CC ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS, CARBOHYDRATE-

CC BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL HYDROPHOBIC

CC PROTEINS (SP-B AND SP-C).

CC -1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.

DR EMBL; M13686; G190670; -.

DR EMBL; K03475; G190672; -.

DR EMBL; M30838; G190565; -.

DR EMBL; M68519; G338049; -.

DR FIR; A24622; LNHUP5.

DR FIR; A25720; LNHUP6.

DR FIR; B25720; LNHUP1.

DR HSSP; P02462; 1BBE.

DR MIM; 178630; -.

DR PROSITE; PS00615; C-TYPE_LECTIN_1; 1.

DR PROSITE; PS0041; C-TYPE_LECTIN_2; 1.

KW GLYCOPROTEIN; CALCIUM; SURFACE FILM; GASEOUS EXCHANGE; HYDROXYLATION;

KW SIGNAL; LECTIN; COLLAGEN; REPEAT; POLYMORPHISM.

FT SIGNAL 1 20

FT CHAIN 21 248 PULMONARY SURFACTANT PROTEIN A.

FT DOMAIN 28 100 COLLAGEN-LIKE (WITH HYDROXYPROLINE).

FT DISULFID 153 248 C-TYPE LECTIN (SHORT FORM).

FT DISULFID 224 238 BY SIMILARITY.

FT CARBOHYD 207 207 PROBABLE.

FT VARIANT 9 9 N -> T.

FT VARIANT 66 66 M -> T.

FT VARIANT 73 73 D -> N.

FT VARIANT 81 81 I -> V.

FT VARIANT 85 85 C -> R.

FT CONFLICT 19 19 A -> V (IN REF. 2).

FT CONFLICT 45 45 D -> H (IN REF. 2).

FT CONFLICT 50 50 L -> V (IN REF. 3).

FT CONFLICT 54 54 P -> L (IN REF. 2).

FT CONFLICT 91 91 P -> A (IN REF. 3).

FT CONFLICT 100 100 P -> R (IN G190565).

FT CONFLICT 247 247 E -> D (IN REF. 3).

SQ SEQUENCE 248 AA; 26214 MW; 56968AA0 CRC32;

Query Match

Best Local Similarity 32.1%; Score 63; DB 1; Length 248;

Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 95 GPPGLPAHLDEELQAT 110

QY 3 GPPGLQRLQRLQAS 18

RESULT 14

ID FREC_SINP2 STANDARD; PRT; 256 AA.

AC F19737;

DT 01-FEB-1991 (REL. 17, CREATED)

DT 01-FEB-1991 (REL. 17, LAST SEQUENCE UPDATE)

DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)

DE HYPOHETICAL 28.7 KD PROTEIN IN RECA 3'REGION.

OS SYNECHOCOCUS SP. (STRAIN PCC 7002) (AGMENELLUM QUADRUPPLICATUM).

OC CYANOBACTERIA; GRACILICUTES; OXYPHOTOBACTERIA;

OC CYANOBACTERIA (BLUE-GREEN ALGAE); CHROCOCCALES.

[1]

SEQUENCE FROM N.A.

RX MEDLINE; 90130334.

RA MURPHY R.C., GASPARICH G.E., BRYANT D.A., PORTER R.D.;

WQREH
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Thu Jul 30 09:12:01 1998; Maspar time 2.79 Seconds
251.630 Million cell updates/sec

Similar output not generated.

Title: >US-08-938-548A-4
Description: (1-28) from US08938548A.pap
Perfect Score: 196
Sequence: 1 RSGPPGLQRLQLRLLQASGNHAAIGLTM 28

Scoring table: PAM 150
Gap 15

Searched: 69111 seqs, 25083644 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swiss1

Statistics: Mean 31.467; Variance 48.238; scale 0.652

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	71	36.2	398	1	VE2_HPV63 REGULATORY PROTEIN E2.	5.14e+01
2	69	35.2	736	1	ALD_MOUSE ADRENOLEUKODYSTROPHY P	1.11e+00
3	67	34.2	79	1	NIFU_FRAAL NIFU PROTEIN.	2.34e+00
4	67	34.2	442	1	HRB8_STRCO RNA POLYMERASE PRINCIP	2.34e+00
5	67	34.2	1477	1	YORL_YEAST OLIGOMYCIN RESISTANCE	2.34e+00
6	66	33.7	461	1	VE2_HPV09 REGULATORY PROTEIN E2.	3.39e+00
7	65	33.2	1692	1	CYAA_SCHPO ADENYLATE CYCLASE (EC	4.88e+00
8	64	32.7	394	1	RT04_YEAST MITOCHONDRIAL 40S RIBO	7.01e+00
9	64	32.7	400	1	ASSV_SYNY3 ARGININOSUCCINATE SYNT	7.01e+00
10	64	32.7	690	1	YTER_ERV PROBABLE DNA PACKAGING	7.01e+00
11	64	32.7	879	1	YDBH_ECOLI HYPOTHETICAL 96.8 KD P	7.01e+00
12	63	32.1	199	1	IL1L_MOUSE INTERLEUKIN-11 PRECURS	1.00e+01
13	63	32.1	248	1	PSPA_HUMAN PULMONARY SURFACTANT-A	1.00e+01
14	63	32.1	256	1	YREC_SYNP2 HYPOTHETICAL 28.7 KD P	1.00e+01
15	63	32.1	324	1	ANNXIII_III (LIPOCORTI	1.00e+01
16	63	32.1	429	1	UL88_HCMVA PROTEIN UL88.	1.00e+01
17	63	32.1	878	1	ECR_DRONE ECDYSONE RECEPTOR (ECD	1.00e+01
18	63	32.1	972	1	POLS_IPNVJ STRUCTURAL POLYPROTEIN	1.00e+01
19	63	32.1	1157	1	PEX1_PICPA PEROXISOME BIOSYNTHESI	1.00e+01
20	63	32.1	1992	1	TR12_HUMAN THYROID RECEPTOR INTER	1.00e+01
21	62	31.6	187	1	MT28_YEAST TRANSCRIPTIONAL ACTIVA	1.43e+01
22	62	31.6	511	1	CTAQ_THEAQ THERMOSTABLE CARBOXYPE	1.43e+01
23	62	31.6	624	1	SIR_SYNP7 SULFITE REDUCTASE (FER	1.43e+01

24	62	31.6	928	1	VGLB_MCMVS GLYCOPROTEIN B PRECURS	1.43e+01
25	62	31.6	972	1	POLS_IPNVN STRUCTURAL POLYPROTEIN	1.43e+01
26	62	31.6	1176	1	NIR_NEUCR NITRITE REDUCTASE (NAD	1.43e+01
27	62	31.6	1690	1	KFIA_HUMAN KINESIN-LIKE PROTEIN K	1.43e+01
28	62	31.6	1695	1	KFIA_MOUSE KINESIN-LIKE PROTEIN K	1.43e+01
29	62	31.6	2504	1	FAS_HUMAN FATTY ACID SYNTHASE (E	1.43e+01
30	61	31.1	180	1	NEF_HV2NZ NEGATIVE FACTOR (F-PRO	2.02e+01
31	61	31.1	249	1	PSPA_PIG PULMONARY SURFACTANT-A	2.02e+01
32	61	31.1	410	1	B7_USTMA MATING-TYPE LOCUS ALLE	2.02e+01
33	61	31.1	520	1	YB77_YEAST HYPOTHETICAL 57.9 KD P	2.02e+01
34	61	31.1	799	1	AFSK_STRCO SERINE/THREONINE PROTE	2.02e+01
35	60	30.6	179	1	SPMB_BACSU SPORE MATURATION PROTE	2.85e+01
36	60	30.6	199	1	IL1L_HUMAN INTERLEUKIN-11 PRECURS	2.85e+01
37	60	30.6	274	1	DCOP_MYCBO OROTIDINE 5'-PHOSPHATE	2.85e+01
38	60	30.6	302	1	YP95_CAEEL HYPOTHETICAL 34.8 KD P	2.85e+01
39	60	30.6	367	1	YOPM_YERPE OUTER MEMBRANE PROTEIN	2.85e+01
40	60	30.6	380	1	F812_MOUSE FACTOR VII INTRON 22	2.85e+01
41	60	30.6	442	1	HRB6_XANCV PROBABLE ATP SYNTHASE	2.85e+01
42	60	30.6	495	1	ALAT_HUMAN ALANINE AMINOTRANSFERA	2.85e+01
43	60	30.6	495	1	ALAT_RAT ALANINE AMINOTRANSFERA	2.85e+01
44	60	30.6	652	1	RPSD_CAUCR RNA POLYMERASE SIGMA F	2.85e+01
45	60	30.6	5217	1	HTS1_COCCA HC-TOXIN SYNTHETASE (E	2.85e+01

ALIGNMENTS

RESULT 1						
ID	VE2_HPV63	STANDARD;	PRT;	398 AA.		
AC	Q07850;					
DT	01-OCT-1994 (REL. 30, CREATED)					
DT	01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)					
DT	01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)					
DE	REGULATORY PROTEIN E2.					
GN	E2.					
OS	HUMAN PAPILLOMAVIRUS TYPE 63.					
OC	VIRIDAE; DS-DNA NONENVELOPED VIRUSES; PAPOVAVIRIDAE; PAPILLOMAVIRUSSES.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE; 93276568.					
RA	EGAWA K., DELIUS H., MATSUKURA T., KAWASHIMA M., DE VILLIERS E.M.;					
RL	VIROLOGY 194:789-799(1993).					
CC	- FUNCTION: E2 IS A TRANSCRIPTIONAL TRANSACTIVATOR CAPABLE OF					
CC	ACTIVATING A CONDITIONAL ENHANCER IN THE VIRAL LONG CONTROL					
CC	REGION (LCR). E2 BINDS TO THE 5'-ACCGNCCGGT-3' PALINDROMIC					
CC	SEQUENCE.					
CC	- SUBUNIT: BINDS DNA AS A DIMER.					
DR	EMBL; X70828; G312096; .					
DR	HSP; P11299; 2BOP.					
KW	EARLY PROTEIN; TRANSCRIPTION REGULATION; ACTIVATOR; DNA-BINDING;					
KW	TRANS-ACTING FACTOR; PHOSPHORYLATION.					
SQ	SEQUENCE 398 AA; 45450 MW; C9BB0CE0 CRC32;					

Query Match 36.2%; Score 71; DB 1; Length 398;
Best Local Similarity 55.6%; Pred. No. 5.14e+01;
Matches 10; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

DB	295	RSPPKGGQSKRLRLQEA	312
QY	1	RSGPPGLQRLQLLQAS	18

RESULT 2						
ID	ALD_MOUSE	STANDARD;	PRT;	736 AA.		
AC	P48410;					
DT	01-FEB-1996 (REL. 33, CREATED)					
DT	01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)					
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)					
DE	ADRENOLEUKODYSTROPHY PROTEIN HOMOLOG (ALDP).					
GN	ALD OR ALDGH.					
OS	MUS MUSCULUS (MOUSE).					
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;					
OC	EUTHERIA; RODENTIA.					
RN	[1]					

CC the repeat lacking two Cys residues. The same motif is found in
 CC complement components C3a, C4a and C5a; the inverse is found in
 CC albumin, vitamin D-binding protein and alpha-fetoprotein. The
 CC disulphide-stabilized loop structure is thought to be conserved.
 CC The type II motif, related to repeats found in epidermal growth
 CC factor precursor is a 6-Cys pattern repeated nine times, although
 CC the ninth repeat in the A form is incomplete. Four of the repeats,
 CC (2, 3, 4 and 9) differ from the EGF-like motif in having a 4-6
 CC residue insertion between cysteines 4 and 5 instead of the usual
 CC single residue. Embodied within repeats 5, 6, 7 and 8 is the
 CC consensus sequence for Asp and Asn hydroxylation. The 7th repeat
 CC contains the consensus O-glycosylation sequence, CXCXPC.
 CC Immediately following each repeat is a pentapeptide with the
 CC consensus sequence XD(I/V)(D/N)E. Fibulin binds to the cyto-
 CC plasmic domain of the beta1 subunit of integrin adhesion
 CC receptors in a cation-dependent, EDTA-reversible manner. It can
 CC be used to manipulate adhesion of cells to fibronectin, collagen,
 CC laminin, and possibly also other proteins. Antibodies reactive
 CC with the protein have important diagnostic and therapeutic uses.
 CC See also R11147, R11148 and R11150.
 CC Sequence 601 AA;

Query Match 10.1%; Score 96; DB 2; Length 601;
 Best Local Similarity 57.1%; Pred. No. 5.91e+00;
 Matches 16; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

Db 10 vplp1lllgglallagvdadvllleacc 37
 QY 13 VTLLLLLLPPALLSLGVDQAQPLPD-CC 39

RESULT 3

ID R11150 standard; Protein; 683 AA.
 AC R11150;
 DC 21-MAY-1991 (first entry)
 DE Fibulin C.
 KW Beta-1 integrin; adhesion; receptor; fibronectin.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT peptide 1..29
 FT protein /label= signal sequence
 FT protein /label= fibulin C
 FT modified_site 98
 FT modified_site /label= N-linked glycosylation
 FT modified_site 535
 FT modified_site /label= N-linked glycosylation
 FT modified_site 339
 FT region /label= N-linked glycosylation
 FT region 36..144
 FT region /label= type I motif
 FT region 36..69
 FT region /label= repeat unit 1
 FT region 112..144
 FT region /label= repeat unit 2
 FT region 144..179
 FT region /label= Glu/Asp-rich region
 FT region 180..566
 FT region /label= type II motif
 FT region 180..214
 FT region /label= repeat unit 1
 FT region 215..219
 FT region /label= consensus pentapeptide
 FT region 220..260
 FT region /label= repeat unit 2
 FT region 261..265
 FT region /label= consensus pentapeptide
 FT region 266..306
 FT region /label= repeat unit 3
 FT region 307..311
 FT region /label= consensus pentapeptide
 FT region 312..354
 FT region /label= repeat unit 4

FT region 355..359
 FT /label= consensus pentapeptide
 FT region 360..397
 FT /label= repeat unit 5
 FT region 398..402
 FT /label= consensus pentapeptide
 FT region 403..439
 FT /label= repeat unit 6
 FT region 440..444
 FT /label= consensus pentapeptide
 FT region 445..479
 FT /label= repeat unit 7
 FT region 480..484
 FT /label= consensus pentapeptide
 FT region 485..529
 FT /label= repeat unit 8
 FT region 530..535
 FT /label= consensus pentapeptide

PN W09102755-A.
 PD 07-MAR-1991.
 PF 17-AUG-1990; U04662.
 PR 18-AUG-1989; US-395773.
 PA (LJOL-) LA JOLLA CANCER RES.
 PA (AMNA-) AMER NAT RED CROSS.
 PI Ruoslahti EI, Argraves WS;
 DR WPI: 91-087250/12.
 DR N-PSDB; Q11010.
 PT Purified fibulin, DNA encoding it and antibodies reactive with it
 PT - useful as diagnostic and therapeutic component.
 PS Claim 10; Fig 5; 56pp; English.
 CC The fibulin A, B and C forms are identical from their N-terminals
 CC to a divergence point at posn. 566 (terminal codon in fibulin A)
 CC after which they are distinct, encoding polypeptides of 566, 601
 CC and 683 residues resp. All three forms are rich in cysteine (11 %)
 CC and analysis wrt no. and spacing of the Cys residues revealed two
 CC types of repeat motif (I and II). The type I motif, CC(X)12C-
 CC (X)9-10C(X)6CC is repeated twice, separated by an imperfect form of
 CC the repeat lacking two Cys residues. The same motif is found in
 CC complement components C3a, C4a and C5a; the inverse is found in
 CC albumin, vitamin D-binding protein and alpha-fetoprotein. The
 CC disulphide-stabilized loop structure is thought to be conserved.
 CC The type II motif, related to repeats found in epidermal growth
 CC factor precursor is a 6-Cys pattern repeated nine times, although
 CC the ninth repeat in the A form is incomplete. Four of the repeats,
 CC (2, 3, 4 and 9) differ from the EGF-like motif in having a 4-6
 CC residue insertion between cysteines 4 and 5 instead of the usual
 CC single residue. Embodied within repeats 5, 6, 7 and 8 is the
 CC consensus sequence for Asp and Asn hydroxylation. The 7th repeat
 CC contains the consensus O-glycosylation sequence, CXCXPC.
 CC Immediately following each repeat is a pentapeptide with the
 CC consensus sequence XD(I/V)(D/N)E. Fibulin binds to the cyto-
 CC plasmic domain of the beta1 subunit of integrin adhesion
 CC receptors in a cation-dependent, EDTA-reversible manner. It can
 CC be used to manipulate adhesion of cells to fibronectin, collagen,
 CC laminin, and possibly also other proteins. Antibodies reactive
 CC with the protein have important diagnostic and therapeutic uses.
 CC See also R11147, R11148 and R11149.
 SQ Sequence 683 AA;

Query Match 10.1%; Score 96; DB 2; Length 683;
 Best Local Similarity 57.1%; Pred. No. 5.91e+00;
 Matches 16; Conservative 3; Mismatches 8; Indels 1; Gaps 1;

Db 10 vplp1lllgglallagvdadvllleacc 37
 QY 13 VTLLLLLLPPALLSLGVDQAQPLPD-CC 39

RESULT 4

ID P70134 standard; Protein; 440 AA.
 AC P70134;
 DT 26-APR-1991 (first entry)
 DE Natural recombinant human lecithin:cholesterol acyltransferase.

```

RESULT 9
AC R50009 standard; Protein; 241 AA.
ID AC
AC R50009;
DE 29-SEP-1994 (first entry)
DE Platelet-derived Growth Factor PDGF-B precursor.
DE Platelet-derived Growth Factor; heterodimer; PDGF-AB;
KW recombinant protein production; PDGF-B chain;
KW bicistronic vector system.
OS Homo sapiens.
FH Key
FH peptide
FT 1..81
FT /label= pre-peptide
FT /label= mature_PDGF-B
FT /note= "mature peptide is encoded by nucleotides
FT 283-609 of Q58725"
FN W09405785-AA.
FN 17-MAR-1994.
FN 26-AUG-1993; DE-2294.
FN 27-AUG-1992; DE-228458.
PA (BEIE ) BEIERSDORF AG.
PA (GBFE ) GBF GES BIOTECH FORSCHUNG GMBH.
PI Achterberg V, Dirks W, Dorschner A, Eichner W, Hauser H;
PI Meyer-Ingold W, Mielke H, Wirth M, Doerschner A;
PI WPI: 94-101190/12.
DR N-PSDB; Q58725.
DR New multicistronic expression units - for producing equimolar
PT amts. of polypeptide(s) in mammalian cells as hosts
PT Claim 11; Page 28; 109pp; German.
PS A PDGF-AB heterodimer is recombinantly produced using a bicistronic
PS expression unit in which a sequence responsible for internal
PS translation start is located between cistrons coding for the PDGF-B
PS and PDGF-A chains. The preferred PDGF-B sequence for inclusion in
PS the bicistronic construct is Q58725 which codes for the precursor
PS amino acid sequence R50009.
SQ Sequence 241 AA;

Query Match 9.3%; Score 89; DB 9; Length 241;
Best Local Similarity 43.6%; Pred. No. 1.94e+01;
Matches 17; Conservative 11; Mismatches 6; Indels 5; Gaps

Db 44 lgrllhgdpgedgaeldlnmrshrsggelesla-rgr 81
|||||: : : : : : : : : : : : : : : : : :
QY 79 LQRLQAN-GNH-AAGI-LTMGR-RAGAELEPYPCGRR 113

RESULT 10
ID R63469 standard; protein; 241 AA.
AC R63469;
DE 28-JUN-1995 (first entry)
DE Recombinant platelet derived growth factor B c-sis.
DE Recombinant platelet derived growth factor B c-sis; rPDGF B c-sis;
KW recombinant platelet derived growth factor B c-sis;
KW wound healing; antibody production.
OS Synthetic.
OS EP-622456-A.
PN PD 02-NOV-1994.
PN EP-622456-A.
PF 10-MAR-1988; 302116.
PF 13-MAR-1987; US-025344.
PR 19-FEB-1988; US-152045.
PA (AMGE-) AMGEN INC.
PA Nicolson MA, Thomason AR;
PI WPI: 94-334642/42.
DR New vectors expressing platelet-derived growth factor B
DR epitope(s) - and contg. c-sis or v-sis genes portions; are used
PT to promote wound healing and produce antibodies
PS Disclosure; Fig 2; 49pp; English.
PS R63469 describes the amino acid sequence of recombinant platelet
PS derived growth factor B c-sis (rPDGF B c-sis). It may be used to
PS promote wound healing, or as an antigen to prepare antisera or
PS monoclonal antibodies specific for epitopes on the B chain of PDGF.
SQ Sequence 241 AA;

```


8D05

U.S. DEPARTMENT OF COMMERCE
Patent and Trademark Office

SEARCH REQUEST FORM

7-781

Requestor's

Name: _____

Serial

Number: _____

Date: _____

Phone: _____

Art Unit: _____

Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

FOR OFFICIAL USE ONLY**STAFF USE ONLY**

Date completed: 8-5-98
Searcher: M. SPENCER
Terminal time: 3
Elapsed time: prep Maude 15
CPU time: _____
Total time: 18
Number of Searches: 1
Number of Databases: 6

Search Site

☒ STIC
☒ CM-1
☐ Pre-S

Type of Search

☒ N.A. Sequence
☒ A.A. Sequence
☐ Structure
☐ Bibliographic

Vendors

☒ to MPSRCH
☐ STN
☐ Dialog
☐ APS
☐ Geninfo
☐ SDC
☐ DARC/Questel
☐ Other